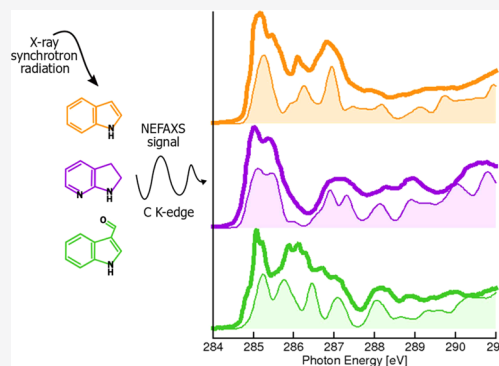


# Carbon and Nitrogen K-Edge NEXAFS Spectra of Indole, 2,3-Dihydro-7-azaindole, and 3-Formylindole

Aurora Ponzi,\* Elisa Bernes, Daniele Toffoli,\* Giovanna Fronzoni, Carlo Callegari, Alessandra Ciavardini, Michele Di Fraia, Robert Richter, Kevin C. Prince, Hanan Sa'adeh, Michele Devetta, Davide Faccialà, Caterina Vozzi, Lorenzo Avaldi, Paola Bolognesi, Mattea Carmen Castrovilli, Daniele Catone, Marcello Coreno, and Oksana Plekan\*

**ABSTRACT:** The near-edge X-ray absorption fine structure (NEXAFS) spectra of indole, 2,3-dihydro-7-azaindole, and 3-formylindole in the gas phase have been measured at the carbon and nitrogen K-edges. The spectral features have been interpreted based on density functional theory (DFT) calculations within the transition potential (TP) scheme, which is accurate enough for a general description of the measured C 1s NEXAFS spectra as well as for the assignment of the most relevant features. For the nitrogen K-edge, the agreement between experimental data and theoretical spectra calculated with TP-DFT was not quite satisfactory. This discrepancy was mainly attributed to the many-body effects associated with the excitation of the core electron, which are better described using the time-dependent density functional theory (TDDFT) with the range-separated hybrid functional CAM-B3LYP. An assignment of the measured N 1s NEXAFS spectral features has been proposed together with a complete description of the observed resonances. Intense transitions from core levels to unoccupied antibonding  $\pi^*$  states as well as several transitions with mixed-valence/Rydberg or pure Rydberg character have been observed in the C and N K-edge spectra of all investigated indoles.

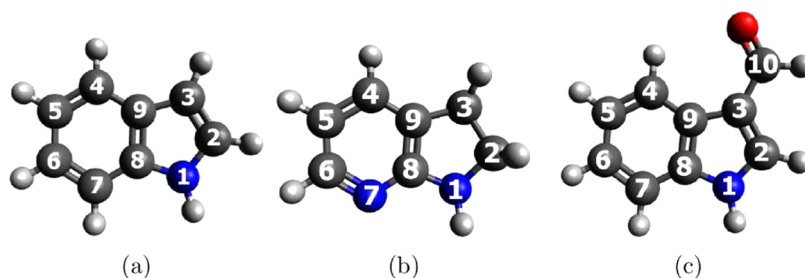


## INTRODUCTION

Heterocycles such as indoles<sup>1</sup> represent a very important class of compounds that play a major role in biology, chemistry, and medicine;<sup>2</sup> for example, indole is the parent molecule of the essential amino acid tryptophan. There has been an increasing interest in the use of indoles as they are important scaffolds in a wide range of synthetic drugs, which are clinically used for anti-inflammatory, analgesic, anticancer, and antidepressant therapeutic applications.<sup>3–5</sup>

Additionally, in recent years, indoles have received growing attention as coating materials. It has been shown that indole-coated surfaces have desirable properties such as selectivity, sensitivity, and stability.<sup>6</sup> Thus, they have widespread applications compared to bare surfaces and can be used as selective electrodes, which are sensitive to various cationic and anionic inorganic species, as a biosensor for biological molecules, or as a protection of metallic surfaces against corrosion.<sup>7,8</sup> Many intrinsic properties (for example, the polymerization mechanism) of bicyclic heterocyclic molecules are masked by their environment or by their interactions with it.<sup>9,10</sup> In this case, gas-phase data can provide a better understanding of the properties of indoles in the absence of perturbations due to their interactions with the surroundings.

Because of their importance, indoles have been extensively studied using various spectroscopic techniques. Recently, the photoinduced structural transformations of selected indoles isolated in cryogenic noble-gas matrices were studied by Reva and co-workers.<sup>11–13</sup> It was found that the tendency of indole and other indole derivatives to undergo hydrogen-atom transfer from N1 to the C3 atom of the indole ring (see Figure 1) seems to be strongly related to the formation of C3-centered radicals, which facilitate the reattachment of the labile hydrogen atom at this position. Early photoelectron spectra (PES) of nitrogen-containing heterocycles were measured with a He I source (21.2 eV) and assigned on the basis of molecular orbital (MO) calculations.<sup>14–16</sup> A comprehensive electronic-structure analysis of boron-containing (BN) indoles and indole was carried out by Chrostowska et al. using a combined UV-PES and computational chemistry approach.<sup>17</sup> Other exper-



**Figure 1.** Schematic structures of (a) indole (I), (b) 2,3-dihydro-7-azaindole (7-AI), and (c) 3-formylindole (3-FI). The C (black) and N (blue) atoms are labeled.

imental studies of the electronic properties of indole and its derivatives have employed optical (including vibrationally and rotationally resolved) methods<sup>18–22</sup> as well as time-resolved ion and photoelectron spectroscopy techniques.<sup>23–26</sup> Theoretical studies of bicyclic molecules such as indole have been mainly focused on geometry optimization and calculations of vibrational spectra using the Hartree–Fock (HF) method, density functional theory (DFT), and MP2 methods.<sup>27–31</sup>

Despite the relevance of indoles, X-ray spectroscopic information about these heterocycles remains limited. Recently, the valence band and core-level spectra of gaseous indole, 2,3-dihydro-7-azaindole, and 3-formylindole were measured by X-ray photoemission spectroscopy and interpreted by quantum-chemistry calculations.<sup>32</sup> An inner-shell photoionization study of indole in the gas phase was published by Kierspel and co-authors,<sup>33</sup> and detailed information about its photoionization and photofragmentation spectra upon single-photon inner-shell ionization was reported. To the best of our knowledge, complete electronic structure data recorded by X-ray photoelectron spectroscopy (XPS) and near-edge X-ray absorption fine structure (NEXAFS) spectroscopy techniques exist only for gaseous 3-methylindole.<sup>34,35</sup>

It is well known that core electron spectroscopies are useful tools to investigate the electronic structure of free molecules, ranging from simple diatomic compounds to very complex ones. This is essentially due to the localized character of the core hole so that the relative intensity of transitions to the unoccupied valence levels can be used as a probe of their spatial distribution and local symmetry.<sup>36</sup> NEXAFS spectroscopy is an ideal method to probe the electronic properties of materials through the absorption fine structure in the spectra, which corresponds to transitions from core orbitals to unoccupied orbitals.<sup>37</sup> Simple rules have been proposed for spectral analysis and applied to a wide range of species, such as the one-center rules, the building block principle, and symmetry-based selection rules.<sup>38,39</sup> However, for a detailed and precise assignment of measured NEXAFS spectra, detailed theoretical calculations are required.

Thus, the present work reports the experimental NEXAFS spectra at the C and N K-edges of gaseous indole (I) and two derivatives: 2,3-dihydro-7-azaindole (7-AI) and 3-formylindole (3-FI), shown in Figure 1. The experimental NEXAFS spectra are supported by quantum-chemistry calculations performed in the framework of the transition potential DFT (TP-DFT) and time-dependent DFT (TDDFT) approaches (see [Theoretical](#) and [Computational Details](#) sections).

## METHODS

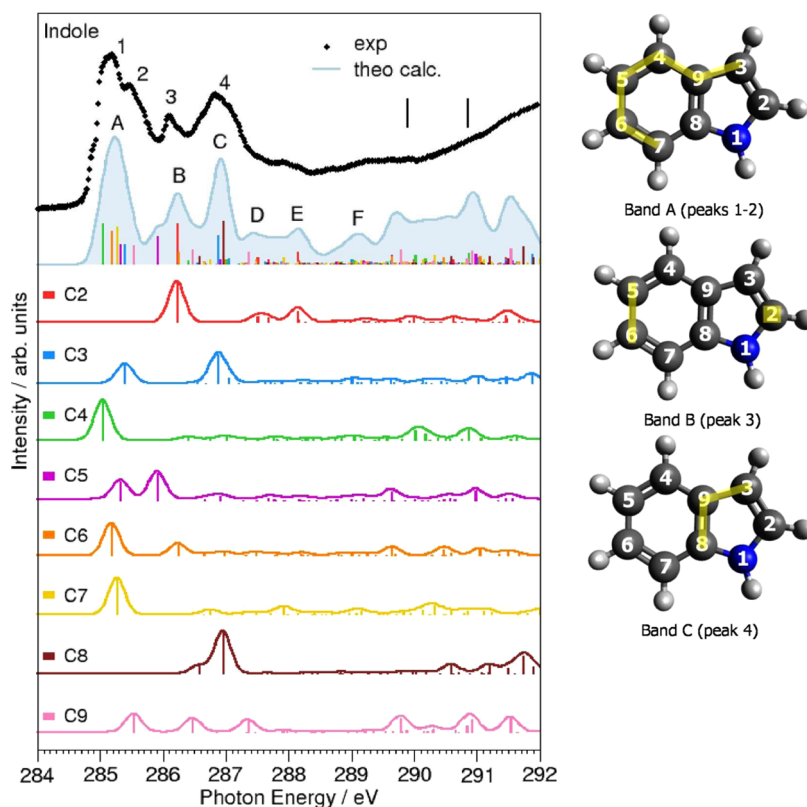
**Experimental Details.** The samples were obtained from Sigma-Aldrich in the form of crystalline powder with a

minimum purity of 99% and used without any further purification. Indole and 7-AI were introduced into the system via an effusive needle at room temperature, while 3-FI was evaporated from a home-built furnace with an effusive nozzle, at a temperature of 350 K. During the experiment, the sample quality was periodically monitored by valence band photoemission and photoionization mass spectroscopies. The energies of peaks in the valence spectra were similar to previously published He I spectra,<sup>15–17</sup> and no effects of thermal decomposition were observed.

The measurements were performed at the Gas-Phase Photoemission and Circular Polarization beamlines at the ELETTRA synchrotron light source in Trieste (Italy).<sup>40,41</sup> The NEXAFS spectra were recorded by collecting the total ion yield signal using a channel electron multiplier placed close to the ionization region. The spectra at the carbon and nitrogen K-edges were taken with a resolution of 50 and 60 meV, respectively. The X-ray absorption spectra were normalized to the photon flux measured simultaneously by a photodiode. In a separate measurement, the photon energy scale was calibrated to known resonances by measuring the spectra of a mixture of the sample and the following gases: CO<sub>2</sub> (C 1s → π\*, 290.77 eV),<sup>42</sup> N<sub>2</sub> (N 1s → π\*, 400.87 eV),<sup>43</sup> and CO<sub>2</sub> (O 1s → π\*, 535.4 eV).<sup>44</sup>

Additionally, the O K-edge spectrum of 3-FI was measured, but due to the water present in the sample and/or the residual gas of the experimental chamber, it was difficult to define a final absorption feature corresponding to this molecule (see [Figure S1](#), Supporting Information (SI)).<sup>45</sup>

**Theoretical Details.** To simulate X-ray absorption spectra, we used the TP-DFT approach that includes most of the relaxation effects upon the formation of the core hole, at an affordable computational cost.<sup>46</sup> Excitation energies are calculated as the energy difference of the final virtual and the core initial TP-MOs involved in the transition. In DFT-TP, the ionization potential values (IPs), based on Koopmans’ theorem,<sup>47</sup> are calculated as the opposite of the TP eigenvalue associated with the initial core orbital. Since this approach generally leads to a less attractive potential, the resulting absolute transition energies are usually too large compared to the experimental ones. To overcome this, one can first compute the IPs within the ΔKS (ΔSCF Kohn–Sham) scheme, allowing a full relaxation of the ionized core hole, and then shift the TP excitation energies by a value corresponding to the difference between the energy of the initial core-excited TP-MOs and IP<sup>ΔKS</sup>, where the IP<sup>ΔKS</sup> is given by the difference between the energy of the *N*-1-electronic configuration and that of the *N*-electron configuration. The energy of the 1s<sup>-1</sup> ionic state is calculated through a Kohn–Sham (KS) spin-polarized unrestricted scheme. A further rigid shift of the



**Figure 2.** Left side: C K-edge NEXAFS spectrum of indole. The experimental spectrum (black diamonds) is shown together with the total theoretical line shape (solid blue line) and the partial  $C_i$  contributions (colored vertical lines). Partial  $C_i$  contributions convoluted with Gaussian profiles with full width at half-maximum (FWHM) = 0.3 eV are presented as colored spectra in the bottom panels. The calculated profile obtained by the TP-DFT/B3LYP potential has been shifted by  $-0.9$  eV to match the first experimental peak. The experimental ionization thresholds are also shown (black vertical solid bars).<sup>32</sup> Right side: for the main peaks appearing in the C 1s NEXAFS spectrum of indole, the C atoms that mainly ( $f \times 10^2 \geq 0.80$ ) contribute to their intensity are highlighted.

theoretical profiles has been applied for a better comparison with the measured spectra. The values of the applied energy shifts are reported in the captions of the figures.

For the simulation of the NEXAFS spectra with TDDFT, we used the Casida formulation of linear-response TDDFT,<sup>48</sup> in which the excitation spectrum is obtained through the solution of the eigenvalue equation using Davidson’s iterative algorithm.<sup>49,50</sup> As core-excitation energies lie very high in the excitation spectrum,<sup>51</sup> some approximations must be adopted, such as the core–valence separation approximation by Cederbaum et al.<sup>52,53</sup>

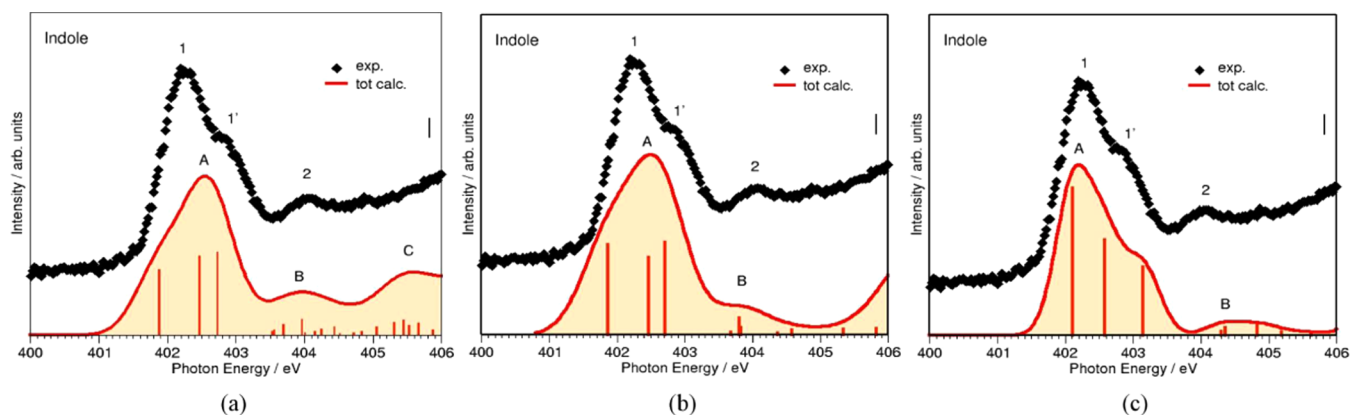
**Computational Details.** The geometry optimizations of the three molecules were performed at the DFT B3LYP/aug-cc-pVTZ level through the Gaussian09 program.<sup>54</sup> The calculations of C, N, and O K-edge NEXAFS spectra were carried out with the Amsterdam Density Functional (ADF) program.<sup>55,56</sup> The TP-DFT<sup>46,57–59</sup> was employed with the generalized gradient approximation (GGA) PW86x Perdew<sup>60</sup> and hybrid B3LYP<sup>61–63</sup> xc potentials, while TDDFT calculations in the linear-response regime<sup>48</sup> used both B3LYP and the range-separated hybrid CAM-B3LYP exchange–correlation (xc) potentials.<sup>64</sup> In the case of 3-formylindole, all calculations were performed for its more stable *1H-trans* isomer.<sup>11</sup>

For the TP-DFT calculations with PW86x and B3LYP, an even-tempered quadruple- $\zeta$  with three polarization and three diffuse functions (ET-QZ3P-3DIFFUSE in the ADF database) Slater-type orbitals (STO) basis set has been employed for the

core-excited C, N, and O atoms, to properly describe transitions to diffuse Rydberg states. A triple  $\zeta$  polarized (TZP in the ADF database) basis set of STOs was adopted for the core orbitals of nonexcited C, N, and O atoms, in particular, a frozen core (FC) TZP. The FC 1s basis set was employed to ensure the localization of the half core hole. In the TDDFT calculations, the total number of basis functions was reduced using a TZP basis set for all atoms.

In the C and N K-edge NEXAFS spectra calculations with the PW86x and B3LYP potentials within the TP-DFT scheme, separate computation of the excitation spectrum of each nonequivalent C/N site was performed, and the partial contributions were summed to yield the total spectrum.

C K-edge NEXAFS spectra (and N K-edge for 7-AI) at the TDDFT level can be computed with two different schemes: either the single excitation space consists only of the excitation from one specific core orbital or the reduced space includes all core orbitals. In the latter case, the whole spectrum is obtained from a single run (termed “coupled” in the SI),<sup>45</sup> while in the former case (termed “uncoupled” in the SI),<sup>45</sup> the number of TDDFT runs to be carried out equals the number of nonequivalent centers, and the total spectrum is obtained by summing the contributions of each atom. Both schemes are used in this work, and the results are reported in the SI (Tables S1, S3, and S5 for the C K-shell spectra of I, 7-AI, and 3-FI, respectively, and Table S4 for the N K-shell of 7-AI).<sup>45</sup> Spectra computed with these two protocols compare well with each other.



**Figure 3.** N K-edge spectrum of indole. The experimental spectrum (black diamonds) is shown together with the calculated results (solid red line and red vertical lines). The calculated profiles obtained by (a) TP-DFT/B3LYP, (b) TDDFT/B3LYP, and (c) TDDFT/CAM-B3LYP have been shifted by  $-0.45$ ,  $+12.80$ , and  $+12.55$  eV, respectively, to match the first experimental peak. The experimental ionization threshold is also shown (black vertical solid bar).<sup>32</sup>

## RESULTS AND DISCUSSION

The C 1s and N 1s experimental and theoretical NEXAFS spectra of I, 7-AI, and 3-FI are presented in Figures 2–8, and the data are summarized in Tables 1–6. In Figures 2–5, 7, and 8, the colored partial contributions of all nonequivalent  $C_i$  and Ni atoms are also highlighted. The complete list of the calculated excitation energies and oscillator strengths for the

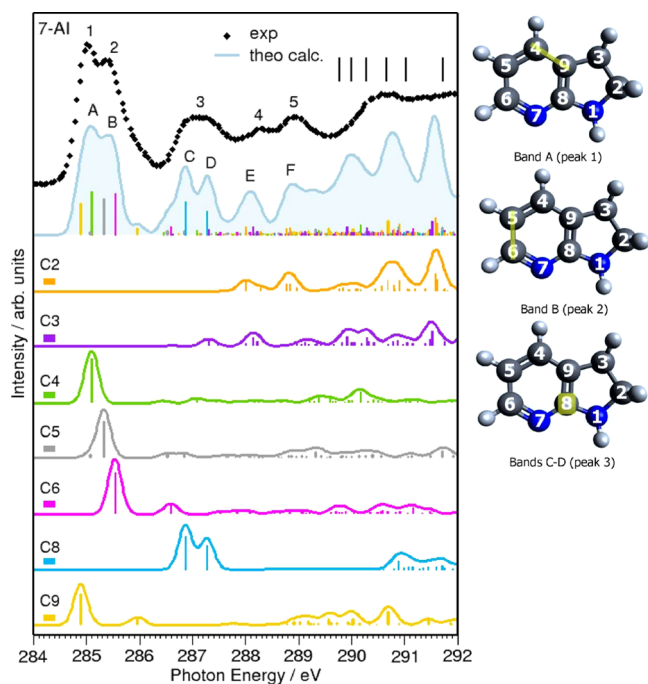
vertical core-level transitions together with the assignment of the corresponding bands of the three heterocycles is available in the SI.<sup>45</sup> In Tables 1–6, only the measured band maxima are shown together with their assignment in terms of the strongest transitions, as predicted by the present calculations.

Here, we present mainly results obtained with TP-DFT combined with the B3LYP potential for both C and N K-edges, as results computed with PW86x do not significantly deviate from them. The calculated C 1s NEXAFS spectra of the compounds match well the experimental data, and they represent a very complex band structure in the 284–292 eV energy range with the maxima labeled by capital letters (see Figures 2, 4, and 7). In the case of indole and 3-FI, for the nitrogen K-edge, the agreement between the experimental results and the theoretical spectral profiles computed with TP-DFT was not quite satisfactory. TDDFT was used to provide a qualitatively correct picture of the measured N 1s NEXAFS spectra, as well as to fully assign the observed features (see the discussion below and Figures 3 and 8). The following discussion is based on the analysis of the experimental and calculated C and N K-edge spectra of the parent molecule indole. The impact of the substitution of one carbon atom in the benzene ring of indole by nitrogen (7-AI) and of the presence of a formyl group (3-FI) has been probed by the measured photoabsorption profiles. Overall, the most evident effect of these structural changes resulted in a shift of some resonances toward higher energy, as well as a redistribution of intensities and the appearance of additional transitions. Throughout the text, TP-MOs refer to the MOs with the core hole on  $C_i$  sites obtained from the TP calculation.

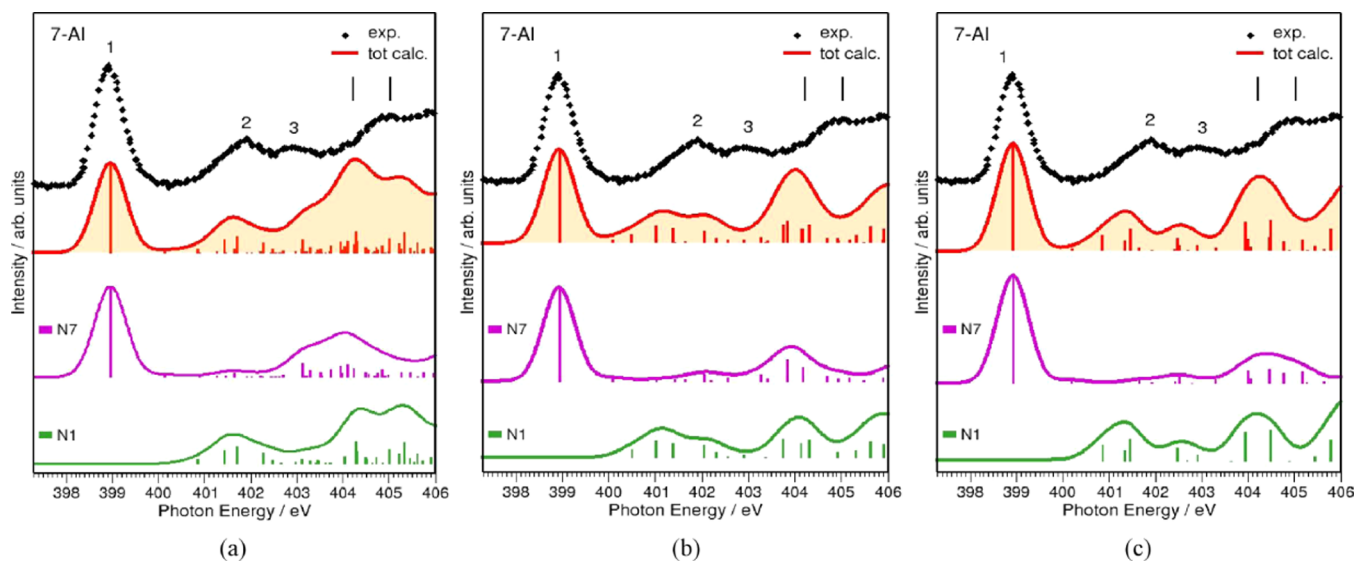
### C 1s and N 1s Photoabsorption Spectra of Indole.

Indole is planar and consists of fused benzene and pyrrole rings (see Figure 1a). It has ten  $\pi$ -electrons, is aromatic, and readily undergoes electrophilic substitution reactions similar to benzene.<sup>65</sup> Indole possesses four unoccupied antibonding MOs of  $\pi$  character corresponding to the three benzene  $\pi^*$  MOs and the  $\pi^*$  MO of the pentacyclic ring.

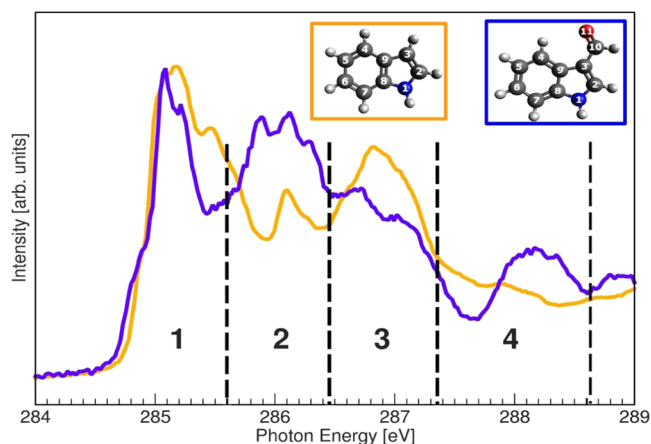
The experimental C K-edge NEXAFS spectrum of indole (see Figure 2, left side) is characterized by a first resonance (peak 1) in the photon energy range between 284.8 and 285.9 eV with a maximum of intensity centered at 285.15 eV and a shoulder (peak 2) at a higher photon energy of 285.5 eV. There are two further resonances labeled as peaks 3 and 4; the



**Figure 4.** Left side: C K-edge NEXAFS spectrum of 2,3-dihydro-7-azaindole. The experimental spectrum (black diamonds) is shown together with the total theoretical line shape (solid blue line) and the partial  $C_i$  contributions (colored vertical lines). Partial  $C_i$  contributions convoluted with Gaussian profiles with FWHM = 0.3 eV are presented as colored spectra in the bottom panels. The calculated profile obtained by TP-DFT/B3LYP has been shifted by  $-1.1$  eV to match the first experimental peak. The experimental ionization thresholds are also shown (black vertical solid bars).<sup>32</sup> Right side: for the main peaks appearing in the C 1s NEXAFS spectrum of 7-AI, the C atoms that mainly ( $f \times 10^2 \geq 0.80$ ) contribute to their intensity are highlighted.



**Figure 5.** N K-edge NEXAFS spectrum of 2,3-dihydro-7-azaindole. The experimental spectrum (black diamonds) is shown together with the calculated results (solid red line and red vertical lines). The calculated profiles obtained by (a) TP-DFT/B3LYP, (b) TDDFT/B3LYP, and (c) TDDFT/CAM-B3LYP have been shifted by  $-0.7$ ,  $+12.30$ , and  $+11.90$  eV, respectively, to match the first experimental peak. The experimental ionization thresholds are also shown (black vertical solid bars).<sup>32</sup>



**Figure 6.** C K-edge experimental NEXAFS spectrum of 3-formylindole (blue) compared to the C K-edge spectrum of indole (yellow). The dashed black lines delimit four different spectral regions (labeled as 1–4).

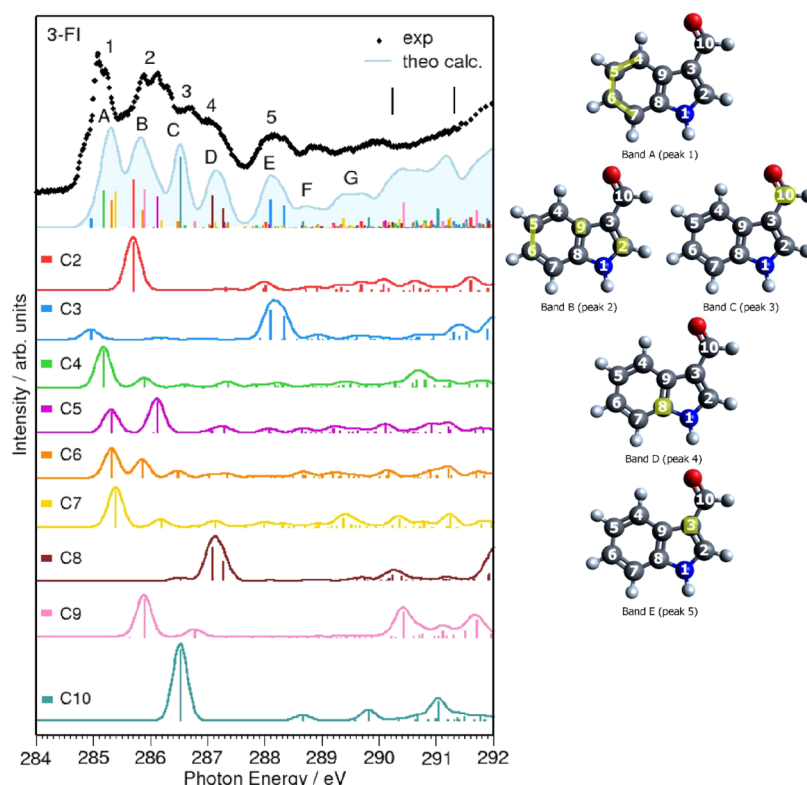
first one is quite sharp, appearing at 286.1 eV, and the second one is broader and more intense than the former, with a maximum at 286.8 eV. In the spectral interval starting from about 288.5 eV, the experimental spectrum is no longer resolved.

The calculated total C 1s spectrum of indole shows a first band (A), in the region of peaks 1–2 of the experimental profile, entirely due to C 1s  $\rightarrow$  lowest unoccupied molecular orbital (LUMO) ( $1\pi^*$ ) transitions from all of the  $C_i$  atoms of the molecule, except for those directly bonded to the heteroatom (i.e., C2 and C8) (see Table 1). A lower excitation energy is predicted for this set of atoms, as their 1s orbitals are more screened compared to C2- and C8 1s orbitals. The dominant contribution results from the excitations to LUMO from C4, C6, and C7, namely, three of the atoms belonging to the benzene ring. The stronger intensity of these transitions to the LUMO can be related to the greater localization of the TP-LUMO on these sites (see Table S8, SI).<sup>45</sup> Based on the

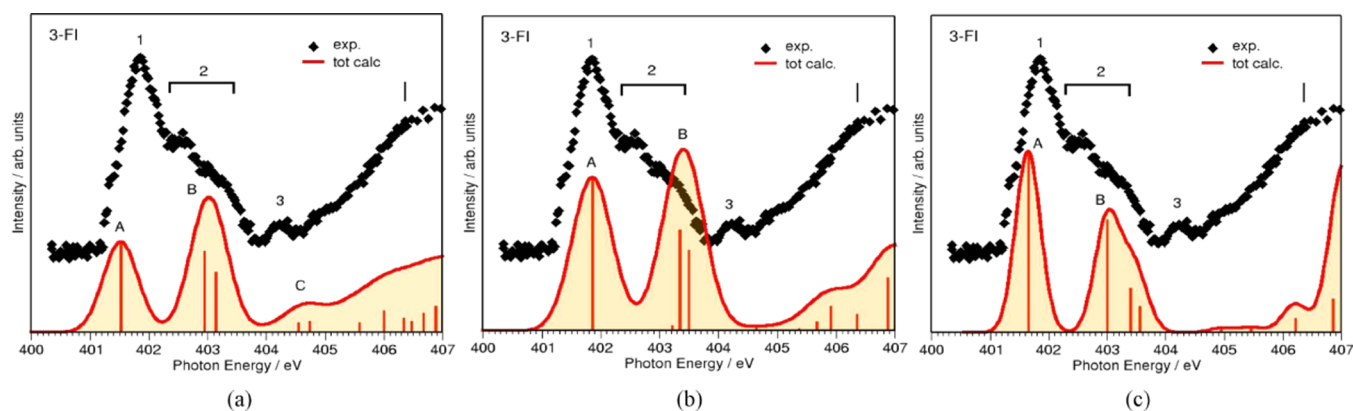
present calculations, the shoulder appearing in the experimental spectrum at 285.5 eV can be assigned to core-excitations from C3, C5, and C9 carbons to  $\pi^*$  MOs. Indeed, in comparison with our previously studied 3-methylindole molecule,<sup>35</sup> the same feature was also assigned to the excitation of carbon atoms within the pyrrole ring.

Higher C 1s  $\rightarrow$  LUMO excitation energies are predicted for C2 and C8, which are directly bonded to the heteroatom and therefore less screened, in agreement with the higher C 1s XPS binding energies (BEs) (experimental values are 290.61 eV for C2 and 290.66 eV for C8).<sup>32</sup> The C2 1s  $\rightarrow$  LUMO ( $1\pi^*$ ) and C8 1s  $\rightarrow 2\pi^*$  excitations, characterized by a comparable value of the oscillator strengths, are mainly responsible for the appearance of bands B (286.1 eV) and C (286.8 eV). By considering the two bands separately, band B is due to transitions to the LUMO ( $1\pi^*$ ) from C2 and to the  $2\pi^*$  orbital from C6 and C5. The calculated profile of the total spectrum shows a shoulder appearing at a lower energy with respect to the main peak of the second band due to the C5 1s  $\rightarrow 2\pi^*$  transition. Band C, although derived from contributions of several excitations, mainly results from the transition to  $3\pi^*$  from C3, i.e., one of the atoms of the pyrrole ring, and from the transition to  $2\pi^*$  from C8, the site directly bonded to the nitrogen. The analysis of the core-excitations from the individual  $C_i$  sites shows that the excitations to the LUMOs are generally more intense than those to the higher energy orbitals; however, C3 and C8 exhibit an opposite behavior. For C8, this can be explained by considering the larger  $C_i$   $2p_z$  AO contribution to the LUMO + 1 ( $2\pi^*$ ) compared to that of the LUMO, whereas for C3, the relaxation effects following the core-hole formation seem to play a dominant role (see Table S9, SI).<sup>45</sup>

Our calculations predict several transitions of both valence and Rydberg characters in the spectral region starting from around 287.5 eV (D–F bands), where the experimental spectrum shows only a weak modulation without any pronounced features. In particular, shoulder D can be assigned to C7 1s  $\rightarrow 4\pi^*/$ Rydberg and C2 1s  $\rightarrow 2\pi^*$  excitations, whereas feature E is essentially due to a mixed-valence ( $3\pi^*$ )-



**Figure 7.** Left side: C K-edge NEXAFS spectrum of 3-formylindole. The experimental spectrum (black diamonds) is shown together with the total theoretical line shape (solid blue line) and the partial  $C_i$  contributions (colored vertical lines). Partial  $C_i$  contributions convoluted with Gaussian profiles with FWHM = 0.3 eV are presented as colored spectra in the bottom panels. The calculated profile obtained by TP-DFT/B3LYP has been shifted by +1.0 eV to match the first experimental peak. The experimental ionization thresholds are also shown (black vertical solid bars; note that the third IP = 292.65 eV is out of the present photon energy range).<sup>32</sup> Right side: for the main peaks appearing in the C 1s NEXAFS spectrum of 3-FI, the C atoms that mainly contribute ( $f \times 10^2 \geq 0.80$ ) to their intensity are highlighted.



**Figure 8.** N K-edge NEXAFS spectrum of 3-formylindole. The experimental spectrum (black diamonds) is shown together with the calculated results (solid red line and red vertical lines). The calculated profiles obtained by (a) TP-DFT/B3LYP, (b) TDDFT/B3LYP, and (c) TDDFT/CAM-B3LYP have been shifted by  $-0.25$ ,  $+13.25$ , and  $+12.55$  eV, respectively, to match the first experimental peak. The experimental ionization threshold is also shown (black vertical solid bar).<sup>32</sup>

Rydberg transition from C2. Before the ionization threshold, core-excitations from all  $C_i$  sites to Rydberg orbitals are responsible for the shape of the spectral envelope.

The N 1s experimental and theoretical NEXAFS spectra of indole are shown in Figure 3, and the energies of the observed resonances are summarized in Table 2. The absorption spectrum is characterized by a strong asymmetric band A (peak 1) between 401.5 and 403.5 eV, with a maximum intensity at around 402.25 eV. This resonance is due to the N 1s core-excitation to the LUMO ( $1\pi^*$ ) (see Table 2). In this

energy range, the TP-DFT calculation predicts three transitions with an intensity distribution that does not correctly reflect the experimental profile. In particular, the lower intensity of the shoulder at 402.82 eV, labeled as 1', compared to peak 1, observed experimentally, is not well reproduced (see Figure 3a). No significant variation in the intensity distribution is obtained using another exchange–correlation potential (i.e., PW86x), so that the reason for the disagreement with the experiment cannot be attributed to the choice of the DFT xc potential. Besides, this shoulder has also been observed in the

**Table 1. Peak Assignments for the C K-Edge NEXAFS Spectrum of Indole<sup>a,b</sup>**

	theory (TP-DFT/B3LYP)			experiment				
	energy (eV) shifted	$f \times 10^2$	site assignment	maximum	peak			
band A	285.04	2.65	C4 LUMO ( $1\pi^*$ )	285.15	1			
	285.18	2.15	C6					
	285.27	2.42	C7					
	285.32	1.28	C5					
	285.39	1.29	C3					
band B	285.53	1.23	C9	285.50	2			
	286.22	2.66	C2 LUMO ( $1\pi^*$ )					
	286.24	0.80	C6 $2\pi^*(C=C)$					
	285.90	1.83	C5 $2\pi^*(C=C)$					
	286.47	0.93	C9 $2\pi^*(C=C)$					
band C	286.58	0.52	C8 LUMO ( $1\pi^*$ )	286.80	4			
	286.75	0.31	C7 $\sigma^*(C-H)$					
	286.87	1.87	C3 $3\pi^*(C=C)$					
	286.91	0.31	C5 $\sigma^*(C-H)$					
	286.95	2.80	C8 $2\pi^*(C=C)$					
	286.97	0.27	C6 $\sigma^*(C-H)$					
	287.00	0.27	C4 $\sigma^*(C-H)$					
	287.04	0.35	C3 $\sigma^*(C-H)$					
	287.35	0.81	C9 $3\pi^*(C=C)$					
	287.48	0.25	C6 $3\pi^*(C=C)$ -Rydberg mixed					
	287.50	0.42	C2 $\sigma^*(C-H)$ -Rydberg mixed					
	band D	287.67	0.30			C2 $2\pi^*(C=C)$	287.90	
		287.92	0.43			C7 $4\pi^*(C=C)$ -Rydberg mixed		
band E	288.14	0.76	C2 $3\pi^*(C=C)$ -Rydberg mixed					
band F	289.00	0.25	C3 Rydberg					

<sup>a</sup>Calculated excitation energies (eV) have been shifted by  $-0.9$  eV to match the first experimental peak. <sup>b</sup>Only transitions with  $f \times 10^2 \geq 0.25$  are reported.

**Table 2. Peak Assignments for the N K-Edge NEXAFS Spectrum of Indole<sup>a,b</sup>**

	theory (TP-DFT/B3LYP)			experiment	
	energy (eV) shifted	$f \times 10^2$	assignment	maximum	peak
band A	401.88	0.59	LUMO ( $1\pi^*$ )	402.25	1
	402.46	0.71	$\sigma^*(N-H)$ -Rydberg	402.82 <sub>(shoulder)</sub>	1'
	402.73	0.74	$2\pi^*(C=C)$		
band B	403.96	0.14	Rydberg	404.10	2
band C	405.31	0.12	Rydberg	404.92	
	405.44	0.14			
	405.66	0.11			

<sup>a</sup>Calculated excitation energies (eV) have been shifted by  $-0.45$  eV to match the first experimental peak. <sup>b</sup>Only transitions with  $f \times 10^2 \geq 0.1$  are reported.

nitrogen K-edge spectrum of 3-methylindole (402.5 eV)<sup>35</sup> and pyrrole (402.6 eV)<sup>66</sup> and assigned to a  $N 1s \rightarrow 3p\sigma/\sigma^*(N-H)$  Rydberg-valence excitation. The strong mixing between the  $3p\sigma$  and the  $\sigma^*(N-H)$  MOs, which is responsible for the rather high intensity of the Rydberg state, was discussed in detail by Duflo et al. in the case of gaseous pyrrole.<sup>66</sup>

We, therefore, investigated the possible impact of the coupling among the  $N 1s^{-1}$  excitations, included in a description based on linear-response TDDFT. The results obtained by TDDFT with the hybrid B3LYP xc kernel slightly improved the intensity distribution (see Figure 3b), still indicating the influence of the configuration mixing (see Table S12, SI).<sup>45</sup> The use of the range-separated hybrid CAM-B3LYP is instead able to fully resolve the discrepancy with the experiment, both for the energy separation and for the intensity distribution of the spectral features (see Figure 3c). The analysis of the TDDFT eigenvectors indicates that CAM-B3LYP accounts for a stronger coupling of singly excited configurations compared to B3LYP (see Table S12, SI).<sup>45</sup> Also, the second intense line, contrary to the other two that can be assigned to transitions into pure  $\pi^*$  orbitals, corresponds to a mixed  $\sigma^*(N-H)$  valence-Rydberg transition. Consistent with previous literature, the Rydberg excitations are well described by the range-separated functionals such as CAM-B3LYP, while the B3LYP functional describes the same excitations less accurately.<sup>67,68</sup> Overall, taking into account the  $1h-1p$  configuration mixing by the TDDFT description together with including more HF exchange for long electron-electron distances by the range separation scheme provides a better description of the  $N 1s$  NEXAFS spectrum of indole compared to the simpler TP-DFT scheme.

The next band B (peak 2, Figure 3), barely visible in the experimental spectrum, appears at 404.1 eV. Based on our calculations, it arises from a group of weak transitions to virtual orbitals with both valence and Rydberg characters. Excitation of these states is consistent with the spectral intensity at 404.6 and 404.0 eV for gaseous pyrrole<sup>66</sup> and 3-methylindole,<sup>35</sup> respectively. The hardly resolved resonance at 404.92 eV (band C) in the  $N 1s$  spectrum of indole was also reported previously for pyrrole at 405.2 eV, and it has been assigned to the  $N 1s \rightarrow \sigma^*(N-H)$  excitation.<sup>66</sup>

**C 1s and N 1s Photoabsorption Spectra of 2,3-Dihydro-7-azaindole.** The substitution of one carbon atom in the benzene ring of indole by nitrogen (see Figure 1b) forms 7-AI and induces differences in its photoabsorption spectra. In contrast to indole, 7-AI is a nonplanar compound due to the  $sp^3$  hybridization of the C2 and C3 atoms in the five-membered ring.

The experimental C 1s NEXAFS spectrum of 7-AI is dominated by two strong peaks (1 and 2) at 285.05 and 285.35 eV, respectively, and shows a few less-pronounced features in the energy range of 286–292 eV (see Figure 4, left side). The first well-resolved maximum, A, is due to a pair of transitions to the LUMO ( $1\pi^*$ ) from C9 and C4. The second maximum B is due to a second pair of transitions to  $2\pi^*$  and LUMO from C5 and C6. All of these C-atoms belong to the six-membered ring. As for band A of indole, the first two peaks appearing in the spectrum of 7-AI are mainly due to the excitations from the  $1s$  orbitals of the more screened atoms. The most evident effect of the change in hybridization of C2 and C3 is in a shift to higher energies of the excitation from C3, which, in this case, contributes with two less intense  $\sigma^*$  and Rydberg transitions to the following bands. This is also reflected in the higher XPS BE value of C3 in passing from indole (289.54 eV) to 7-AI (290.80 eV),<sup>32</sup> with a difference in the absolute value of 1.26 eV greater than that found for all other C-atoms.

By comparing the C 1s NEXAFS spectra of indole and 7-AI, one can observe that peak 3 in the spectrum of indole, mainly due to the transition to LUMO from C2, does not have a

corresponding peak in the spectrum of 7-AI, as remarked above due to the lack of planarity in the five-membered ring. The peak at 287.2 eV is mainly contributed by the transition from the C8 atom, directly bonded to both N1 and N7 atoms (see Figure 1b). In particular, our calculations show that the wide resonance corresponding to bands C and D originates from the envelope of two distinct transitions, namely, C8 1s  $\rightarrow$   $1\pi^*$  and C8 1s  $\rightarrow$   $2\pi^*$  (see Table 3). Peak 3 is slightly shifted

**Table 3. Peak Assignments for the C K-Edge NEXAFS Spectrum of 2,3-Dihydro-7-azaindole<sup>a,b</sup>**

	theory (TP-DFT/B3LYP)				experiment	
	energy (eV) shifted	$f \times 10^2$	site	assignment	maximum	peak
band A	284.90	2.21	C9	LUMO ( $1\pi^*$ )	285.05	1
	285.10	3.11	C4	LUMO ( $1\pi^*$ )		
band B	285.33	2.57	C5	$2\pi^*(C=C)$	285.35	2
	285.54	2.96	C6	LUMO ( $1\pi^*$ )		
band C	285.97	0.43	C9	$2\pi^*(C=C)$	287.20	3
	287.08	0.31	C4	$\sigma^*(C-H)$		
	287.30	0.40	C3	$\sigma^*(C-H)$		
	286.53	0.25	C5	Rydberg		
	286.59	0.58	C6	$2\pi^*(C=C)$		
	286.88	2.39	C8	LUMO ( $1\pi^*$ )		
band D	287.28	1.68	C8	$2\pi^*(C=C)$	287.20	3
band E	288.00	0.57	C2	$\sigma^*(C-H)$	288.30	4
	288.14	0.54	C3	Rydberg		
	288.28	0.25	C2	$2\pi^*(C=C)$		
band F	288.77	0.44	C2	Rydberg	288.95	5
	288.85	0.48	C2	Rydberg		
	289.19	0.26	C9	Rydberg		
	289.32	0.40	C5	Rydberg		
	289.38	0.31	C4	Rydberg		
	289.58	0.42	C9	Rydberg		

<sup>a</sup>Calculated excitation energies (eV) have been shifted by  $-1.1$  eV to match the first experimental peak. <sup>b</sup>Only transitions with  $f \times 10^2 \geq 0.25$  are reported.

to a higher energy with respect to the analogous peak in the C 1s NEXAFS spectrum of indole. This shift can be explained by considering that C8 in 7-AI is less screened than the corresponding C8 in indole and a larger excitation energy is then predicted for it.

Finally, bands E and F are assigned to transitions from C2 and C3 1s core-level atoms belonging to the five-membered ring: C2 1s  $\rightarrow$   $\sigma^*(C-H)$  and C3 1s  $\rightarrow$  Rydberg.

In Figure 5, the experimental and theoretical N K-edge NEXAFS spectra of 7-AI are reported and compared. The N 1s NEXAFS spectrum is dominated by peak 1 at 398.93 eV, which shows a slightly asymmetric line shape with a tail on the high-energy side. This asymmetry is due to an unresolved vibrational fine structure, caused by vibrational excitations generated by the change in the molecular potential upon excitation. Based on our calculations, the first band is due to the core-excitation from N7 to the LUMO ( $1\pi^*$ ). The significant localization of the LUMO on the N7 atom gives rise to the higher intensity of the N7 1s  $\rightarrow$   $1\pi^*$  transition compared to the N1 1s  $\rightarrow$   $1\pi^*$ , which has a negligible oscillator strength. This resonance has the same nature as the

strongest peak observed in the N K-edge spectrum of pyridine at around 399 eV.<sup>69–71</sup> All other transitions involving the N atom belonging to the benzene ring are related to the excitations to MOs with a Rydberg character (see Table 4).

**Table 4. Peak Assignments for the N K-Edge NEXAFS Spectrum of 2,3-Dihydro-7-azaindole<sup>a,b</sup>**

theory (TP-DFT/B3LYP)				experiment	
energy (eV) shifted	$f \times 10^2$	site	assignment	maximum	peak
398.96	2.15	N7	LUMO ( $1\pi^*$ )	398.93	1
401.42	0.32	N1	$2\pi^*(C=C)$	401.92	2
401.70	0.39	N1			
402.27	0.25	N1		402.98	3
403.13	0.33	N7	$\sigma^*(C-H)/$		
403.95	0.26	N7	Rydberg		
404.10	0.29	N7			

<sup>a</sup>Calculated excitation energies (eV) have been shifted by  $-0.7$  eV to match the first experimental peak. <sup>b</sup>Only transitions with  $f \times 10^2 \geq 0.25$  are reported.

The experimentally observed feature at 401.92 eV together with the shoulder at 402.98 eV is very similar to that previously described for the N 1s photoabsorption spectrum of indole. In comparison with this, peaks 2 and 3 have been assigned to the transitions N1 1s  $\rightarrow$  LUMO ( $1\pi^*$ ) and N1 1s  $\rightarrow$   $2\pi^*$ , respectively. The region just below the ionization threshold (XPS IPs at 404.23 and 405.03 eV)<sup>32</sup> contains several weak transitions of mixed Rydberg and valence character.

It should be pointed out that also in the case of 7-AI, TP-DFT combined with the B3LYP potential cannot provide a quantitative description of the N 1s NEXAFS spectrum for this molecule, especially for the photon energy range above the first intense experimental peak 1 (see Figure 5a). The TDDFT results, both with B3LYP and CAM-B3LYP, are instead in very good agreement with the experimental spectrum in the whole energy region and correctly describe the relative intensities of the experimentally verified peaks 2 and 3 (Figure 5b,c and Table 4).

**C 1s and N 1s Photoabsorption Spectra of 3-Formylindole.** Indole and 3-formylindole differ by the formyl group bonded to C3 in the pentagonal ring (see Figure 1c). The latter compound has both *1H-trans* and *1H-cis* isomers, which are detectable,<sup>11</sup> and may influence the shape and peak positions of the measured photoabsorption spectrum.

For a visual comparison of the absorption profiles of I and 3-FI, we have reported their experimental C 1s NEXAFS spectra in Figure 6. One can clearly distinguish four different regions (labeled as 1–4 in Figure 6): three of them exhibit spectral shapes with different features for both molecules, while the last one shows a well-defined band only in the case of 3-FI.

The first region (reported in Figure 7) of the experimental C 1s NEXAFS spectrum of 3-FI shows two maxima at 285.08 and 285.20 eV, which, based on the larger oscillator strength values, have been assigned to the LUMO ( $1\pi^*$ ) transitions from the C4–C7 sites, as visible from Figure 7. A minor contribution at the lower energy side of the peaks arises from the C3 to LUMO transition. If we compare the nature of band A with that of the corresponding band in the C 1s NEXAFS spectrum of indole (region 1 in Figure 6), the only difference lies in the lack of the excitation from the C9 atom (see Tables 2 and 5). Indeed, the presence of the formyl group in 3-FI is responsible



**Table 5. Peak Assignments for the C K-Edge NEXAFS Spectrum of 3-Formylindole<sup>a,b</sup>**

	theory (TP-DFT/B3LYP)				experiment	
	energy (eV) shifted	$f \times 10^2$	site	assignment	maximum	peak
band A	284.86	0.55	C3	LUMO	285.08	1
	285.08	2.20	C4	( $1\pi^*$ )		
	285.21	1.27	C5			
	285.22	1.61	C6			
	285.29	2.17	C7			
band B	285.60	2.87	C2	LUMO	285.90	2
				( $1\pi^*$ )		
	285.76	1.03	C6	$2\pi^*$		
	285.79	2.28	C9	LUMO		
				( $1\pi^*$ )		
	285.80	0.52	C4	$2\pi^*$		
	286.01	1.87	C5	$3\pi^*$		
286.09	0.41	C7	$3\pi^*(C=O)$	286.12		
286.38	0.39	C6	$3\pi^*$			
band C	286.43	4.19	C10	LUMO	286.65	3
				( $1\pi^*$ )		
band D	286.67	0.39	C9	$3\pi^*(C=C)$	287.02	4
	286.98	1.95	C8	$2\pi^*(C=C)$		
	287.03	0.34	C7	$\sigma^*(C-H)$		
	287.17	1.16	C8	$3\pi^*(C=C)$		
	287.18	0.32	C5	$\sigma^*(C-H)/$ Rydberg		
	287.24	0.26	C6	$\sigma^*(C-H)$		
band E	287.26	0.32	C4	$\sigma^*(C-H)$	288.19	5
	287.89	0.26	C7	Rydberg		
band E and G	287.91	0.31	C2	$\sigma^*(C-H)/$ Rydberg	288.90	
	288.00	1.71	C3	$\pi^*/$ Rydberg		
	288.23	1.35	C3	$\pi^*/$ Rydberg		
	288.57	0.30	C10	$2\pi^*(C=C)$		
	288.83	0.23	C3	Rydberg		
	289.09	0.19	C6	$4\pi^*$		
	289.10	0.26	C5	$\pi^*$		
	289.27	0.54	C7	$4\pi^*$		
	289.59	0.39	C2	$4\pi^*/$ Rydberg		
	289.72	0.57	C10	Rydberg		
289.98	0.37	C2	$\pi^*/$ Rydberg	290.00		
290.01	0.43	C5	Rydberg			
290.06	0.31	C6	Rydberg			
290.13	0.30	C8	$4\pi^*/$ Rydberg			
	290.18	0.25	C4	Rydberg		

<sup>a</sup>Calculated excitation energies (eV) have been shifted by +1.0 eV to match the first experimental peak. <sup>b</sup>Only transitions with  $f \times 10^2 \geq 0.25$  are reported.

for a shift toward higher energies of the C9 1s  $\rightarrow$  LUMO excitation. This shift is also in good agreement with the higher C 1s binding energy value of the C9 carbon in 3-FI compared to that of indole (290.24 vs 289.76 eV).<sup>32</sup> Furthermore, the stronger intensity of the transitions to the LUMO from C4 and C7 compared to the other two C 1s excitations of band A (i.e., from C5 and C6) can be explained by considering the larger localization of the TP-LUMO on these two sites (see Table S10, SI).<sup>45</sup>

The photon energy region between 285.5 and 286.5 eV (peak 2) contains several transitions of  $\pi$  character as reported

in Figure 7 and Table 5. The nature of band B is similar to that of the corresponding band visible in the C 1s NEXAFS spectrum of indole and originated by the C2 1s  $\rightarrow$  LUMO and C5 1s  $\rightarrow$  LUMO transitions. In addition, a further contribution of the C9 1s  $\rightarrow$  LUMO excitation, appearing in this spectral region (region 2 in Figure 6) because of the presence of the formyl group, is found. In terms of intensity, the dominant contribution is due to the core-excitation from the carbon directly bonded to the heteroatom (C2), as a result of the higher localization of the TP-LUMO on this carbon site (see Table S10, SI).<sup>45</sup>

Region 3 of Figure 6 falls between 286.5 and 287.5 eV in the C 1s NEXAFS spectrum of 3-FI and corresponds to the spectral interval where the rather intense fourth band for indole was seen (see Figure 6). In this region, there are two different bands (C and D) predicted by our calculations, consistent with the two experimental peaks 3 and 4 (see Table 5). As clearly seen from the calculated spectrum, band C is originated from the C10 1s  $\rightarrow$  LUMO transition (see Figure 7).

For both molecules, the observed spectral feature around 287 eV (C and D, respectively, for indole and 3-FI) is due to the excitation from C8. This is also in agreement with the high values of the calculated binding energies for C8 in both I and 3-FI, which are 290.66 and 291.12 eV, respectively.<sup>32</sup> More specifically, for 3-FI, two core-excitations C8 1s  $\rightarrow$   $2\pi^*$  and C8 1s  $\rightarrow$   $3\pi^*$  contribute to peak 4 (Figure 7) together with several low-intensity transitions of valence character (see Table 5). We do not observe the transition from C8 to the LUMO because of the negligible C 2p<sub>z</sub> contribution to the LUMO at this site, unlike the  $2\pi^*$  and  $3\pi^*$  (see Table S11, SI).<sup>45</sup>

The most evident difference that one can observe between the two spectral profiles in region 3 (see Figure 6) lies in their intensity, which is greater in the case of indole. This can be explained by considering the presence of an additional intense transition from the C3 atom, almost degenerate with the transition from C8, contrary to the C 1s NEXAFS spectrum of 3-FI. For the latter, the transition from C3 is shifted toward the higher energy range, giving rise to the next wide band E (see peak 5 in Figure 7), which is absent in the C 1s absorption spectrum of indole (see region 4 in Figure 6). According to the present calculations, several valence and Rydberg transitions with small oscillator strength values also fall below the ionization threshold.

The N K-edge experimental and theoretical NEXAFS spectra of 3-FI are shown in Figure 8. The measured spectrum of 3-FI is characterized by an intense resonance at 401.86 eV due to the core-excitation to the LUMO ( $1\pi^*$ ). The broad and asymmetric shoulder (see region 2 in Figure 8) contains two peaks in its envelope centered at 402.57 and 403.00 eV, respectively. As already discussed for indole (see section above), TP-DFT predicts an intensity distribution that does not match well the experimental profile (Figure 8a), whereas TDDFT combined with CAM-B3LYP leads to a significant variation of both the energy splitting between the transitions and the intensity distribution (Figure 8c). This is due to the inclusion of a strong mixing between 1h-1p configurations (see Table S13, SI)<sup>45</sup> and to a more appropriate description of transitions with contributions of Rydberg MOs. Indeed, the shoulder present in the experimental N 1s NEXAFS spectrum of 3-FI originates from a set of both N 1s  $\rightarrow$   $\pi^*$  and N 1s  $\rightarrow$   $\sigma^*/$ Rydberg transitions (see band B in Table 6). On the other hand, TDDFT/B3LYP predicts only a slight change in the

intensity without an appreciable improvement of the spectrum compared to the TP-DFT (Figure 8b).

**Table 6. Peak Assignments for the N K-Edge NEXAFS Spectrum of 3-Formylindole<sup>a,b</sup>**

	theory (TP-DFT/B3LYP)			experiment	
	energy (eV) shifted	$f \times 10^2$	assignment	maximum	peak
band A	401.80	0.89	LUMO ( $1\pi^*$ )	401.86	1
band B	403.22	0.80	$\sigma^*(\text{N-H})$ -Rydberg	402.57	2
	403.41	0.60	$3\pi^*$	403.00	
band C	404.85	0.10	Rydberg	404.26	3
	404.98	0.11	Rydberg		
	405.83	0.10	$4\pi^*(\text{C=C})$ / Rydberg	405.00	
	406.25	0.21	Rydberg		

<sup>a</sup>Calculated excitation energies (eV) have been shifted by  $-0.25$  eV to match the first experimental peak. <sup>b</sup>Only transitions with  $f \times 10^2 \geq 0.10$  are reported.

The weak resonance at 404.26 eV in the nitrogen K-edge spectrum of 3-FI corresponds to band B observed for indole (see Figures 3 and 8). According to our calculations as well as to results from pyrrole<sup>66</sup> and 3-methylindole,<sup>35</sup> the nature of the current peak is due to a group of weak transitions to virtual orbitals with both valence and Rydberg character.

## CONCLUSIONS

The high-resolution NEXAFS spectra at the C 1s and N 1s ionization thresholds of three biologically important heterocycles, namely, indole, 2,3-dihydro-7-azaindole, and 3-formylindole, have been studied both experimentally and theoretically. The experimental and theoretical spectra at the C and N K-edges show a strong  $\pi^*$  band deriving from transitions to LUMOs. The remaining resonances observed below the ionization threshold are assigned to transitions that involve final states with a mixed-valence/Rydberg or pure Rydberg character.

It has been shown that the present TP-DFT calculations are accurate enough for a general description of the measured C 1s NEXAFS spectra as well as for the assignment of the most relevant features of the studied indoles. All three heterocyclic compounds show a pronounced double structure in the C 1s NEXAFS spectra at about 285 eV, associated with transitions to the orbitals of benzene and pyridine aromatic subunits. The comparison of the measured and calculated carbon K-edge spectra of indole and 3-formylindole strongly suggests that the presence of an extra formyl group is responsible for the shift of some resonances toward higher energy as well as for the appearance of the additional intense transitions belonging to this group.

For the nitrogen K-edge of all investigated indoles, the agreement of the experimental spectra with theoretical profiles computed by the TP-DFT method was quite poor. TDDFT, taking into account the mixing of singly excited configurations that is not covered by TP-DFT, provides a correct picture of the measured N 1s NEXAFS spectra for these molecules. The structure of all observed resonances was correctly predicted by the calculations, and the nature of all transitions was fully assigned. Overall, the N 1s photoabsorption spectra of indole and 3-formylindole are quite similar, as expected from the similar chemical environment of the N atom, while the

spectrum of 2,3-dihydro-7-azaindole shows a rather different structure due to the extra nitrogen atom located in the pyridine ring.

## ASSOCIATED CONTENT

O 1s NEXAFS spectrum of 3-FI (Figure S1); complete list of the calculated excitation energies and oscillator strengths for the vertical core-level transitions together with the assignment of the corresponding bands of the three heterocycles summarized in Tables S1–S7; composition of selected TP-MOs with a core hole on Ci sites in Tables S8–S11; contribution of the single-particle-hole excitations calculated with TDDFT B3LYP/CAM-B3LYP in Tables S12 and S13 (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

**Aurora Ponzi** – Ruđer Bošković Institute, 10000 Zagreb, Croatia; [orcid.org/0000-0001-9095-4366](https://orcid.org/0000-0001-9095-4366); Email: [aponzi@irb.hr](mailto:aponzi@irb.hr)

**Daniele Toffoli** – Dipartimento di Scienze Chimiche e Farmaceutiche, Università degli Studi di Trieste, I-34127 Trieste, Italy; [orcid.org/0000-0002-8225-6119](https://orcid.org/0000-0002-8225-6119); Email: [toffoli@units.it](mailto:toffoli@units.it)

**Oksana Plekan** – Elettra-Sincrotrone Trieste S.C.p.A., 34149 Trieste, Italy; [orcid.org/0000-0002-4692-7018](https://orcid.org/0000-0002-4692-7018); Email: [oksana.plekan@elettra.eu](mailto:oksana.plekan@elettra.eu)

### Authors

**Elisa Bernes** – Dipartimento di Scienze Chimiche e Farmaceutiche, Università degli Studi di Trieste, I-34127 Trieste, Italy

**Giovanna Fronzoni** – Dipartimento di Scienze Chimiche e Farmaceutiche, Università degli Studi di Trieste, I-34127 Trieste, Italy; [orcid.org/0000-0002-5722-2355](https://orcid.org/0000-0002-5722-2355)

**Carlo Callegari** – Elettra-Sincrotrone Trieste S.C.p.A., 34149 Trieste, Italy

**Alessandra Ciavardini** – CERIC-ERIC, 34149 Trieste, Italy

**Michele Di Fraia** – Elettra-Sincrotrone Trieste S.C.p.A., 34149 Trieste, Italy

**Robert Richter** – Elettra-Sincrotrone Trieste S.C.p.A., 34149 Trieste, Italy

**Kevin C. Prince** – Elettra-Sincrotrone Trieste S.C.p.A., 34149 Trieste, Italy; [orcid.org/0000-0002-5416-7354](https://orcid.org/0000-0002-5416-7354)

**Hanan Sa'adeh** – Elettra-Sincrotrone Trieste S.C.p.A., 34149 Trieste, Italy; Department of Physics, The University of Jordan, Amman 11942, Jordan; [orcid.org/0000-0002-2568-5507](https://orcid.org/0000-0002-2568-5507)

**Michele Devetta** – CNR-Istituto di Fotonica e Nanotecnologie (CNR-IFN), 20133 Milano, Italy

**Davide Faccialà** – CNR-Istituto di Fotonica e Nanotecnologie (CNR-IFN), 20133 Milano, Italy

**Caterina Vozzi** – CNR-Istituto di Fotonica e Nanotecnologie (CNR-IFN), 20133 Milano, Italy

**Lorenzo Avaldi** – Istituto di Struttura della Materia-CNR (ISM-CNR), 00133 Rome, Italy; [orcid.org/0000-0002-2990-7330](https://orcid.org/0000-0002-2990-7330)

Paola Bolognesi – Istituto di Struttura della Materia-CNR (ISM-CNR), 00133 Rome, Italy; [orcid.org/0000-0002-6543-6628](https://orcid.org/0000-0002-6543-6628)

Mattea Carmen Castrovilli – Istituto di Struttura della Materia-CNR (ISM-CNR), 00133 Rome, Italy

Daniele Catone – Istituto di Struttura della Materia-CNR (ISM-CNR), 00133 Rome, Italy; [orcid.org/0000-0002-7649-2756](https://orcid.org/0000-0002-7649-2756)

Marcello Coreno – Istituto di Struttura della Materia-CNR (ISM-CNR), 00133 Rome, Italy

## Author Contributions

All authors contributed equally to the data acquisition and treatment. The theory was by E.B., D.T., and A.P. The paper was drafted by O.P. and A.P., refined in consultation with K.C.P. and D.T., and then circulated to all authors who contributed with comments and constructive criticism.

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We gratefully acknowledge the assistance of our colleagues at Elettra for providing good-quality synchrotron light. This research was supported by the Croatian Science Foundation under grant no. HRZZ IP-2016-06-1142. This work was carried out during scientific leave granted to H.S. from the University of Jordan, Amman, Jordan, during the academic year 2018–2019. H.S. acknowledges the TRIL fellowship awarded by the Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, Italy. M.D., D.F., and C.V. acknowledge support from CNR Laboratorio Congiunto “ATTOBIO” and the Italian Ministry of Research and Education with the projects ELI and EUROFEL ESFRI Roadmap. Computational research was supported by Finanziamento per Ricerca di Ateneo, FRA 2015 and FRA 2016, of the Università degli Studi di Trieste. Finally, the COST Action CA18222 Attochem supported by COST (European Cooperation in Science and Technology) is acknowledged.

## ABBREVIATIONS USED

NEXAFS, near-edge X-ray absorption fine structure  
DFT, density functional theory  
TDDFT, time-dependent density functional theory  
TP, transition potential  
TP-DFT, transition potential density functional theory  
GS, ground state  
I, indole  
7-AI, 2,3-dihydro-7-azaindole  
3-FI, 3-formylindole  
SI, supporting information  
MOs, molecular orbitals  
IPs, ionization potential values  
KS, Kohn–Sham  
ALDA, adiabatic local density approximation  
ADF, Amsterdam density functional  
LUMO, lowest unoccupied molecular orbital ( $f \times 10^2$ ), oscillator strength  
xc, exchange and correlation  
BEs, binding energies

## REFERENCES

- (1) Sundberg, R. J. *The Chemistry of Indoles*; Academic Press, New York, 1996.
- (2) Sharma, V.; Kumar, P.; Pathak, D. Biological Importance of the Indole Nucleus in Recent Years: A Comprehensive Review. *J. Heterocycl. Chem.* **2010**, *47*, 491–502.
- (3) Zhang, L. S.; Davies, S. S. Microbial Metabolism of Dietary Components to Bioactive Metabolites: Opportunities for New Therapeutic Interventions. *Genome Med.* **2016**, *8*, No. 46.
- (4) Sravanthi, T. V.; Manju, S. L. Indoles - A Promising Scaffold for Drug Development. *Eur. J. Pharm. Sci.* **2016**, *91*, 1–10.
- (5) Kaushik, N. K.; Kaushik, N.; Attri, P.; Kumar, N.; Kim, C. H.; Verma, A. K.; Choi, E. H. Biomedical Importance of Indoles. *Molecules* **2013**, *18*, 6620–6662.
- (6) Deletioğlu, D.; Hasdemir, E.; Solak, A. O.; Üstündağ, Z.; Güzel, R. Preparation and Characterization of Poly (indole-3-carboxaldehyde) Film at the Glassy Carbon Surface. *Thin Solid Films* **2010**, *519*, 784–789.
- (7) Roman, G.; Pappas, A. Ch.; Demertzi, D. K.; Prodromidis, M. I. Preparation of a 2-(4-fluorophenyl)indole-modified Xerogel and its Use for the Fabrication of Screen-Printed Electrodes for the Electroanalytical Determination of Sulfide. *Anal. Chim. Acta* **2004**, *523*, 201–207.
- (8) Lin, X.; Li, Y. Monolayer Covalent Modification of 5-Hydroxytryptophan on Glassy Carbon Electrodes for Simultaneous Determination of Uric Acid and Ascorbic Acid. *Electrochim. Acta* **2006**, *51*, 5794–5801.
- (9) Udum, Y. A.; Düdükçü, M.; Köleli, F. Electrochemical Polymerization and Spectroscopic Investigation of 2-Methylindole. *React. Funct. Polym.* **2008**, *68*, 861–867.
- (10) Talbi, H.; Monard, G.; Loos, M.; Billaud, D. Theoretical Study of Indole Polymerization. *J. Mol. Struct.: THEOCHEM* **1998**, *434*, 129–134.
- (11) Reva, I.; Lapinski, L.; Lopes Jesus, A. J.; Nowak, M. J. Photoinduced Transformations of Indole and 3-Formylindole Monomers Isolated in Low-Temperature Matrices. *J. Chem. Phys.* **2017**, *147*, No. 194304.
- (12) Lopes Jesus, A. J.; Rosado, M. T. S.; Fausto, R.; Reva, I. UV-Induced Radical Formation and Isomerization of 4-Methoxyindole and 5-Methoxyindole. *Phys. Chem. Chem. Phys.* **2020**, *22*, 22943–22955.
- (13) Nowak, N. J.; Reva, I.; Rostkowska, H.; Lapinski, L. UV-Induced Hydrogen-Atom Transfer and Hydrogen-Atom Detachment in Monomeric 7-Azaindole Isolated in Ar and n-H<sub>2</sub> Matrices. *Phys. Chem. Chem. Phys.* **2017**, *19*, 11447–11454.
- (14) Eland, J. H. D. Photoelectron Spectra of Conjugated Hydrocarbons and Heteromolecules. *Int. J. Mass Spectrom. Ion Phys.* **1969**, *2*, 471–484.
- (15) Domel-Smith, L. N.; Munchausen, L. L.; Houk, K. N. Photoelectron Spectra of Psychotropic Drugs. I. Phenethylamines, Tryptamines, and LSD. *J. Am. Chem. Soc.* **1977**, *99*, 4311–4321.
- (16) Kovač, B.; Klasinc, L.; Stanovnik, B.; Tišler, M. Photoelectron Spectroscopy of Heterocycles. Azaindenes and Azaindolizines. *J. Heterocycl. Chem.* **1980**, *17*, 689–694.
- (17) Chrostowska, A.; Xu, S.; Mazière, A.; Boknevit, K.; Li, B.; Abbey, E. R.; Dargelos, A.; Graciaa, A.; Liu, S.-Y. UV-Photoelectron Spectroscopy of BN Indoles: Experimental and Computational Electronic Structure Analysis. *J. Am. Chem. Soc.* **2014**, *136*, 11813–11820.
- (18) Philips, L. A.; Levy, D. H. The Rotationally Resolved Electronic Spectrum of Indole in the Gas Phase. *J. Chem. Phys.* **1986**, *85*, 1327–1332.
- (19) Berden, G.; Meerts, W. L.; Jalviste, E. Rotationally Resolved Ultraviolet Spectroscopy of Indole, Indazole, and Benzimidazole: Inertial Axis Reorientation in the S<sub>1</sub>(<sup>1</sup>L<sub>b</sub>)←S<sub>0</sub> Transitions. *J. Chem. Phys.* **1995**, *103*, 9596–9606.
- (20) Brand, C.; Küpper, J.; Pratt, D. W.; Meerts, W. L.; Krüger, D.; Tatchen, J.; Schmitt, M. Vibronic Coupling in Indole: I. Theoretical

- Description of the  $^1L_a-^1L_b$  Interaction and the Electronic Spectrum. *Phys. Chem. Chem. Phys.* **2010**, *12*, 4968–4979.
- (21) Küpper, J.; Pratt, D. W.; Meerts, L.; Brand, C.; Tatchen, J.; Schmitt, M. Vibronic Coupling in Indole: II. Investigation of the  $^1L_a-^1L_b$  Interaction Using Rotationally Resolved Electronic Spectroscopy. *Phys. Chem. Chem. Phys.* **2010**, *12*, 4980–4988.
- (22) Nesvadba, R.; Studecký, T.; Uhlíková, T.; Urban, Š. Microwave Spectrum and Molecular Constants of Indole. *J. Mol. Spectrosc.* **2017**, *339*, 6–11.
- (23) Montero, R.; Conde, A. P.; Ovejas, V.; Castaño, F.; Longarte, A. Ultrafast Photophysics of the Isolated Indole Molecule. *J. Phys. Chem. A* **2012**, *116*, 2698–2703.
- (24) Livingstone, R.; Schalk, O.; Boguslavskiy, A. E.; Wu, G.; Bergendahl, L. T.; Stolow, A.; Paterson, M. J.; Townsend, D. Following the Excited State Relaxation Dynamics of Indole and 5-Hydroxyindole Using Time-Resolved Photoelectron Spectroscopy. *J. Chem. Phys.* **2011**, *135*, No. 194307.
- (25) Godfrey, T. J.; Yu, H.; Biddle, M. S.; Ullrich, S. A Wavelength Dependent Investigation of the Indole Photophysics via Ionization and Fragmentation Pump-Probe Spectroscopies. *Phys. Chem. Chem. Phys.* **2015**, *17*, 25197–25205.
- (26) Godfrey, T. J.; Yu, H.; Ullrich, S. Investigation of Electronically Excited Indole Relaxation Dynamics via Photoionization and Fragmentation Pump-Probe Spectroscopy. *J. Chem. Phys.* **2014**, *141*, No. 044314.
- (27) Callis, P. R.; Vivian, J. T.; Slater, L. S. Ab Initio Calculations of Vibrational Spectra for Indole. *Chem. Phys. Lett.* **1995**, *244*, 53–58.
- (28) Sundaraganesan, N.; Umamaheswari, H.; Joshua, B. D.; Meganathan, C.; Ramalingam, M. J. Molecular Structure and Vibrational Spectra of Indole and 5-Aminoindole by Density Functional Theory and Ab Initio Hartree-Fock Calculations. *J. Mol. Struct.: THEOCHEM* **2008**, *850*, 84–93.
- (29) Catalán, J.; de Paz, J. L. G. The Molecular Geometry of Indole. *J. Mol. Struct.: THEOCHEM* **1997**, *401*, 189–192.
- (30) Smith, B. J.; Liu, R. A. A Theoretical Investigation of Indole Tautomers. *J. Mol. Struct.: THEOCHEM* **1999**, *491*, 211–222.
- (31) Caminati, W.; Di Bernardo, S. Microwave Spectrum and Amino Hydrogen Location in Indole. *J. Mol. Struct.* **1990**, *240*, 253–262.
- (32) Plekan, O.; Sa'adeh, H.; Ciavardini, A.; Callegari, C.; Cautero, G.; Dri, C.; Di Fraia, M.; Prince, K. C.; Richter, R.; Sergio, R.; et al. A. Experimental and Theoretical Photoemission Study of Indole and Its Derivatives in the Gas Phase. *J. Phys. Chem. A* **2020**, *124*, 4115–4127.
- (33) Kierspel, T.; Bomme, C.; Di Fraia, M.; Wiese, J.; Anielski, D.; Bari, D.; Boll, R.; Erk, B. M.; Kienitz, J. S.; Müller, N. L. M.; et al. Photophysics of Indole upon X-Ray Absorption. *Phys. Chem. Chem. Phys.* **2018**, *20*, 20205–20216.
- (34) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; Prince, K. C. Valence Photoionization and Photofragmentation of Aromatic Amino Acids. *Mol. Phys.* **2008**, *106*, 1143–1153.
- (35) Zhang, W.; Carravetta, V.; Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; Prince, K. C. Electronic Structure of Aromatic Amino Acids Studied by Soft X-Ray Spectroscopy. *J. Chem. Phys.* **2009**, *131*, No. 035103.
- (36) Hitchcock, A. P. Near Edge Core Excitation Spectra of Complex Molecules. *J. Phys. Colloq.* **1986**, *47*, C8-575–C8-578.
- (37) Stöhr, J. *NEXAFS Spectroscopy*; Springer Series in Surface Sciences; Springer, New York, 1992; Vol. 25.
- (38) Ueda, K. High-Resolution Inner-Shell Spectroscopies of Free Atoms and Molecules Using Soft-X-ray Beamlines at the Third-Generation Synchrotron Radiation Sources. *J. Phys. B: At., Mol. Opt. Phys.* **2003**, *36*, R1–R47.
- (39) Svanberg, S. *Atomic and Molecular Spectroscopy: Basic Aspects and Practical Applications*; Springer, Berlin, 2003.
- (40) Prince, K. C.; Blyth, R. R.; Delaunay, R.; Zitnik, M.; Krempasky, J.; Slezak, J.; Camilloni, R.; Avaldi, L.; Coreno, M.; Stefani, G.; et al. The High Resolution Gas Phase Photoemission Beamline, Elettra. *SJ. Synchrotron Radiat.* **1998**, *5*, 565–568.
- (41) Derossi, A.; Lama, F.; Piacentini, M.; Prosperi, T.; Zema, N. High Flux and High Resolution Beamline for Elliptically Polarized Radiation in the Vacuum Ultraviolet and Soft X-Ray Regions. *Rev. Sci. Instrum.* **1995**, *66*, 1718–1720.
- (42) Tronc, M.; King, G. C.; Read, F. H. Carbon K-shell Excitation in Small Molecules by High-Resolution Electron Impact. *J. Phys. B: At. Mol. Phys.* **1979**, *12*, 137–157.
- (43) Sodhi, R. N. S.; Brion, C. E. Reference Energies for Inner Shell Electron Energy-Loss Spectroscopy. *J. Electron Spectrosc. Relat. Phenom.* **1984**, *34*, 363–372.
- (44) Wight, G. R.; Brion, C. E. K-Shell Energy Loss Spectra of 2.5 keV Electrons in CO<sub>2</sub> and N<sub>2</sub>O. *J. Electron Spectrosc. Relat. Phenom.* **1974**, *3*, 191.
- (45) Supporting Information.
- (46) Triguero, L.; Petterson, L. G. M.; Ågren, H. Calculations of Near-Edge X-Ray-Absorption Spectra of Gas-Phase and Chemisorbed Molecules by Means of Density-Functional and Transition-Potential Theory. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *58*, 8097–8110.
- (47) Koopmans, T. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms. *Physica* **1934**, *1*, 104–113.
- (48) Casida, M. E. Time-Dependent Density Functional Response Theory for Molecules. In *Recent Advances in Density Functional Methods, Part I*; World Scientific, 1995; pp 155–192.
- (49) Davidson, E. R. The Iterative Calculation of a Few of the Lowest Eigenvalues and Corresponding Eigenvectors of Large Real-Symmetric Matrices. *J. Comput. Phys.* **1975**, *17*, 87–94.
- (50) Gross, E. K. U.; Kohn, W. Local Density-Functional Theory of Frequency-Dependent Linear Response. *Phys. Rev. Lett.* **1985**, *55*, 2850–2852.
- (51) Declewa, P.; Fronzoni, G.; Lisini, A.; Stener, M. Molecular Orbital Description of Core Excitation Spectra in Transition Metal Compounds. An Ab-Initio CI Calculation on TiCl<sub>4</sub> and Isoelectronic Molecules. *Chem. Phys.* **1994**, *186*, 1–16.
- (52) Cederbaum, L. S.; Domcke, W.; Schirmer, J. Many-Body Theory of Core Holes. *Phys. Rev. A* **1980**, *22*, 206–222.
- (53) Barth, A.; Cederbaum, L. S. Many-Body Theory of Core-Valence Excitations. *Phys. Rev. A* **1981**, *23*, 1038–1061.
- (54) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. *Gaussian 09*, revision A.02, Gaussian, Inc., Wallingford CT, 2016.
- (55) Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. Towards an Order-N DFT Method. *Theor. Chem. Acc.* **1998**, *99*, 391–403.
- (56) Baerends, E. J.; Ellis, D. E.; Ros, P. Self-Consistent Molecular Hartree-Fock-Slater Calculations I. The Computational Procedure. *Chem. Phys.* **1973**, *2*, 41–51.
- (57) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press, New York, 1989.
- (58) Slater, J. C. Statistical Exchange-Correlation in the Self-Consistent Field. In *Advances in Quantum Chemistry*; Elsevier, 1972; Vol. 6, pp 1–92.
- (59) Ehlert, C.; Gühr, M.; Saalfrank, P. An Efficient First Principles Method for Molecular Pump-Probe NEXAFS spectra: Application to Thymine and Azobenzene. *J. Chem. Phys.* **2018**, *149*, No. 144112.
- (60) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *33*, 8822–8824.
- (61) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (62) Lee, C.; Yang, W.; Parr, R. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *G. Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.
- (63) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem. A* **1994**, *98*, 11623–11627.

- (64) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange–Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (65) Lakhdar, S.; Westermaier, M.; Terrier, F.; Goumont, R.; Boubaker, T.; Ofial, A. R.; Mayr, H. Nucleophilic Reactivities of Indoles. *J. Org. Chem.* **2006**, *71*, 9088–9095.
- (66) Duflot, D.; Hannay, C.; Flament, J. P.; Hubin-Franskin, M. J. Electronic Excitation of Gaseous Pyrrole and Pyrazole by Inner-Shell Electron Energy Loss Spectroscopy. *J. Chem. Phys.* **1998**, *109*, 5308–5318.
- (67) Li, S. L.; Truhlar, D. G. Improving Rydberg Excitations within Time-Dependent Density Functional Theory with Generalized Gradient Approximations: The Exchange-Enhancement-for-Large-Gradient Scheme. *J. Chem. Theory Comput.* **2015**, *11*, 3123–3130.
- (68) Peach, M. J. G.; Cohen, A. J.; Tozer, D. J. Influence of Coulomb-Attenuation on Exchange–Correlation Functional Quality. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4543–4549.
- (69) Baiardi, A.; Mendolicchio, M.; Barone, V.; Fronzoni, G.; Jimenez, G. A. C.; Stener, M.; Grazioli, C.; de Simone, M.; Coreno, M. Vibrationally Resolved NEXAFS at C and N K-Edges of Pyridine, 2-Fluoropyridine and 2,6-Difluoropyridine: a Combined Experimental and Theoretical Assessment. *J. Chem. Phys.* **2015**, *143*, No. 204102.
- (70) Kolczewski, C.; Püttner, R.; Plashkevych, O.; Ågren, H.; Staemmler, V.; Martins, M.; Snell, G.; Schlachter, A. S.; Sant’Anna, M.; Kaindl, G.; et al. Detailed Study of Pyridine at the C 1s and N 1s Ionization Thresholds: the Influence of the Vibrational Fine Structure. *J. Chem. Phys.* **2001**, *115*, 6426–6437.
- (71) Vall-Llosera, G.; Gao, G.; Kivimäki, A.; Coreno, M.; Álvarez Ruiz, J.; de Simone, M.; Ågren, H.; Rachlew, E. The C 1s and N 1s Near Edge X-Ray Absorption Fine Structure Spectra of Five Azabenzene in the Gas Phase. *J. Chem. Phys.* **2008**, *128*, No. 044316.