



**UNIVERSITÀ DEGLI STUDI DI TRIESTE**  
**XXXIII CICLO DEL DOTTORATO DI RICERCA IN**  
**CHIMICA**

**Simulation of core-electron spectroscopies of**  
**gas-phase molecules and adsorbates**  
**by DFT and TDDFT methods**

Settore scientifico-disciplinare: **CHIM/02**

**DOTTORANDA**  
**DOTT. SSA ELISA BERNES**

**COORDINATORE**  
**PROF. ENZO ALESSIO**

**SUPERVISORE DI TESI**  
**PROF. DANIELE TOFFOLI**

**ANNO ACCADEMICO 2019/2020**

# CONTENTS

Overview of the thesis .....	3
Riassunto della tesi .....	6
1. Core-level spectroscopy .....	9
1.1 X-Ray Absorption Spectroscopy (XAS) .....	9
1.1.1 The X-Ray Absorption Cross Section .....	9
1.1.2 Dipole selection rules .....	11
1.1.3 The X-ray Absorption Process .....	12
1.1.4 The fate and nature of core-excited states .....	15
1.1.5 NEXAFS spectroscopy of adsorbed molecules .....	18
1.2 X-Ray Photoelectron Spectroscopy (XPS) .....	24
2. Theoretical methods .....	26
2.1 Hartree-Fock method .....	26
2.1.2 Koopmans' theorem .....	29
2.2 Configuration Interaction (CI) .....	32
2.3 Multi-Configuration Self-Consistent Field (MCSCF) .....	33
2.4 Complete Active Space Self-Consistent Field (CASSCF) .....	34
2.5 Density Functional Theory (DFT) .....	35
2.5.1 The electronic Hamiltonian .....	36
2.5.2 Functionals .....	37
2.5.3 The Thomas-Fermi model .....	37
2.5.4 The birth of modern DFT: Hohenberg and Kohn theorems .....	38
2.5.5 The Kohn-Sham approach .....	41
2.5.6 Exchange and correlation potentials .....	43
2.5.6.1 Local Density Approximation (LDA) .....	44
2.5.6.2 General Gradient Approximation (GGA) .....	45
2.5.6.3 Hybrid exchange and correlation potentials .....	47
2.5.7 The calculation of NEXAFS spectra through DFT: the Transition Potential (TP) method .....	47
2.6 Density of states (DOS) calculations .....	50
2.7 Time-Dependent Density Functional Theory (TDDFT) .....	52
2.7.1 Response Theory (RT) .....	53

2.7.2 Random Phase Approximation (RPA) and TDDFT .....	59
2.8 Theoretical methods for core-electron excitation computations .....	65
2.9 Computational details .....	66
Bibliography .....	69

## Overview of the thesis

The present PhD thesis is focused on the simulation and interpretation of *X-Ray Photoelectron* (XP) and *Near Edge X-ray Absorption Fine Structure* (NEXAFS) spectra of molecules in gas phase and adsorbed on metal surfaces. Core-electron spectroscopies allow to obtain detailed information on the local geometric and chemical environment of the absorbing atom, due to the highly localized nature of the core-hole<sup>1</sup>. NEXAFS spectroscopy is also useful to determine the orientation and geometry of molecules adsorbed on surfaces, giving an atomistic rationalization of factors controlling both the molecular assembly at surfaces and changes in electronic structure following the adsorption process<sup>1</sup>. The main collaborations implicated in these studies have involved experimentalists afferent to the Gas Phase Photoemission and ALOISA beam lines of the Elettra Synchrotron in Trieste, the Department of Physics and Astronomy, Molecular and Condensed Matter Physics of Uppsala University (Sweden), and the Molecular Solids Group of the Philipps University of Marburg (Germany). To establish and quantify the relationship between the measured spectral features and the electronic structure information, XPS and NEXAFS experimental measurements have been supported by theoretical calculations by means of density functional theory (DFT) and its time-dependent generalization (TDDFT) in the linear response regime, which both represent, in conjunction with core-electron spectroscopies, a well-established tool for the electronic structure characterization of materials<sup>2,3</sup>. In many instances, the simulation of the measured XP and NEXAFS spectra has provided a sound assignment of the experimental spectral features as well as a detailed insight on the nature of the virtual states implicated in the X-ray absorption process.

The first systems investigated have been the indole molecule and its 2,3-dihydro-7-azaindole (7-AI) and 3-formylindole (3-FI) derivatives. The choice of these systems has been determined by the fact that indole derivatives play a crucial role in bacterial physiology, ecological balance, and human health<sup>4</sup>. These three heterocycles also find potential applications in chemistry as coating materials, selective electrodes and biosensors<sup>5,6</sup>. The electronic structure of these systems has been investigated through a combined experimental, *Outer Valence Green's function* (OVGF) study of valence band (VB) and C, N, and O 1s XP spectra, in collaboration with the Gas Phase Photoemission beam line of the Elettra Synchrotron in Trieste, and the Ruđer Bošković Institute of Zagreb (Croatia).

A large part of the research activity has dealt with the computation of XP and NEXAFS spectra of a series of gas phase  $\pi$ -conjugated molecular complexes of rising interest in the field of organic molecular electronics<sup>7-9</sup>, in collaboration with the Elettra Synchrotron of Trieste and the Department of Physics and Astronomy, Molecular and Condensed Matter Physics of Uppsala University. The investigated material has been the ambipolar 2,8-bis(diphenylphosphoryl)-dibenzo[b,d]thiophene (PPT), exploited as host of blue phosphorescent organic light-emitting diodes (PhOLEDs) thanks to its wide bandgap and high triplet energy level<sup>9-11</sup>; it is composed by an electron-rich dibenzothiophene (DBT) core, which promotes the hole transport, and two electron-withdrawing diphenylphosphoryl (dPPO) side arms. A first investigation has involved the small thiophene precursor and its benzo-annulated derivatives, namely, benzo[b]thiophene (BBT) and DBT, with the aim of describing the electronic structure of the single electron-rich thiophene and the modification determined by the presence of the benzo-annulated rings in both BBT and DBT. Further studies have been devoted to PPT and its building block moieties, namely, DBT and triphenylphosphine oxide (TPPO), in order to evaluate their impact on the electronic structure of the whole PPT molecular system; this investigation was also expected to show the  $\pi$ -conjugation breaking between the central DBT core and the peripheral dPPO counterparts, which makes the PPT molecular system an interesting ambipolar material, maintaining its characteristic triplet energy level<sup>12</sup>. Another objective was to provide a clearer perspective on the structure of the P–O bond, since its description is still controversial<sup>13</sup>. The XP and NEXAFS spectra have been simulated at the C,O *K-edge* and S,P *L<sub>II,III</sub>-edges* through DFT by using the Transition Potential (TP) method, and TDDFT within the ZORA approximation, respectively.

The research activity has then focused on a metal phtalocyanine (Pc), namely, the titanyl phtalocyanine (TiOPc), a widely-used organic *p*-type semiconductor and photoconductor which also finds promising application in the development of high performance OLEDs and solar cells<sup>14-17</sup>. The aim was the determination of the nature of the TiOPc interaction with the Ag(111) surface through angle-resolved NEXAFS spectra at C, N, and O *K-edge*, since the orientation of the molecule with respect to the surface is still debated<sup>18-20</sup>. This study has been carried out in collaboration with the Molecular Solids Group of the Philipps University of Marburg.

Afterwards, the research activity has been devoted to a series of gas phase boronic systems, namely, boric acid [B(OH)<sub>3</sub>], phenyl boronic acid (PBA), and 1,4-phenyl diboronic acid (PDBA), important building blocks of Covalent Organic frameworks (COFs)<sup>21-23</sup>. This investigation has been motivated by the results of a recent study<sup>24</sup> which hinted at the importance of correlation effects in B1s core-excited state of boronic acid derivatives. Hence, B *K-edge* NEXAFS spectra of B(OH)<sub>3</sub>,

PBA, and PDBA have been calculated through different computational schemes of increasing complexity to evaluate the influence of electron correlation effects; in particular, the purely one-electron scheme of the DFT-TP approach<sup>25</sup>, TDDFT<sup>26</sup>, and the explicitly correlated wave function multiconfigurational self-consistent field (MCSCF)<sup>27,28</sup> method have been employed. Theoretical spectra have been then compared with the experimental ones performed at the ALOISA beam line of the Elettra Synchrotron in Trieste.

Further studies have been devoted to the investigation of the interaction between a series of boronic networks and the Au(111) surface, in collaboration with experimentalists working at the Elettra Synchrotron in Trieste. The modeling of the molecule/metal interface has been facilitated by Scanning Tunnelling Microscopy (STM) images followed by a preliminar optimization of the surface and the adsorbed molecules through a periodic slab methodology within the DFT formalism<sup>29</sup>. Hence, suitable finite size clusters have been cut out for the calculation of B *K-edge* NEXAFS spectra, since it is well known that the finite size cluster model can be efficiently applied to the representation of events localized on the surface or within the bulk of solid systems. Furthermore, it has proven to be very accurate to simulate NEXAFS spectra<sup>30-32</sup>. Hence, the nature of the adsorbate-substrate interaction has been discussed in terms of the differences between the spectral profile of the gas phase molecule and that on the molecule adsorbed on the surface.

The present PhD thesis is divided into two parts, of which the first one (Chapters 1-2) gives a general overview of the theoretical concepts, *i.e.*, the main aspects of NEXAFS spectroscopy, regarding molecules in gas phase and adsorbed on metal surfaces, and the salient steps of DFT and TDDFT, with a special emphasis on the treatment of relaxation effects upon core excitations. The second part of the thesis collects selected research outputs of the three-year period of PhD which have been introduced above. Conclusions and perspectives on the research activity will be finally provided.

## Riassunto della tesi

La presente Tesi di Dottorato ha per oggetto la simulazione e l'interpretazione di spettri XPS (*X-Ray Photoelectron Spectroscopy*) e NEXAFS (*Near Edge X-ray Absorption Fine Structure*) di molecole in fase gassosa e adsorbite su superfici metalliche. Le spettroscopie dei livelli di core consentono di ottenere informazioni dettagliate sull'intorno chimico dell'atomo assorbitore e sulla natura degli orbitali virtuali del sistema investigato, dal momento che gli elettroni di core sono ben localizzati sull'atomo e risentono della simmetria e dell'intensità del campo generato dagli atomi contigui<sup>1</sup>. La spettroscopia NEXAFS è anche utile alla determinazione dell'orientazione e della geometria di molecole adsorbite su superfici, fornendo una razionalizzazione atomistica dei fattori che regolano sia l'assemblaggio su superfici sia i cambiamenti in termini di struttura elettronica in seguito al processo di adsorbimento<sup>1</sup>. Le principali collaborazioni implicate in questi studi hanno coinvolto ricercatori sperimentali afferenti alle *beam lines* Gas Phase Photoemission ed ALOISA del Sincrotrone Elettra di Trieste, al Dipartimento di Fisica ed Astronomia, Fisica della Materia Molecolare e Condensata dell'Università di Uppsala (Svezia), al gruppo di Solidi Molecolari dell'Università Philipps di Marburgo (Germania). Allo scopo di stabilire e quantificare la relazione tra le strutture spettrali misurate e l'informazione sulla struttura elettronica, le misure sperimentali XPS e NEXAFS sono state supportate da calcoli teorici effettuati per mezzo della Teoria del Funzionale Densità (DFT) e la sua generalizzazione dipendente dal tempo (TDDFT) nel regime di risposta lineare, le quali, affiancate alle spettroscopie dei livelli di core, rappresentano uno strumento ben consolidato per la caratterizzazione della struttura elettronica dei materiali<sup>2,3</sup>. In molti casi, la simulazione degli XPS e NEXAFS spettri misurati ha fornito una solida assegnazione delle strutture spettrali sperimentali, nonché una conoscenza dettagliata della natura degli stati virtuali implicati nel meccanismo di assorbimento dei raggi X.

I primi sistemi investigati sono stati la molecola di indolo e due suoi derivati, il 2,3-diidro-7-azaindolo (7-AI) e 3-formilindolo (3-FI). La scelta dei sistemi è stata determinata dal fatto che i derivati dell'indolo rivestono un ruolo fondamentale nella fisiologia batterica, nel bilancio ecologico e nella salute umana<sup>4</sup>. Questi tre eterocicli trovano anche potenziali applicazioni nella chimica come materiali di rivestimento, elettrodi selettivi e biosensori<sup>5,6</sup>. La struttura elettronica di questi sistemi è stata investigata attraverso uno studio sperimentale e OVGf (*Outer Valence Green's Function*) della banda di valenza (VB) e C, N, e O 1s XPS, in collaborazione con la *Gas Phase Photoemission beam line* del Sincrotrone Elettra di Trieste, e l'Istituto Ruđer Bošković di Zagabria (Croazia).

Una larga parte dell'attività di ricerca ha riguardato il calcolo degli spettri XPS e NEXAFS di una serie di complessi molecolari  $\pi$ -coniugati in fase gassosa di crescente interesse nel campo dell'elettronica molecolare organica<sup>7-9</sup>, in collaborazione con il Sincrotrone Elettra di Trieste ed il Dipartimento di Fisica ed Astronomia, Fisica della Materia molecolare e condensata dell'Università di Uppsala. Il materiale investigato è stato l'ambipolare 2,8-bis(difenilfosforil)-dibenzo[b,d]tiofene (PPT), sfruttato come *host* di diodi organici emettitori di luce blu fosforescenti (PhOLEDs) grazie al suo ampio *bandgap* e all'alta energia del livello dello stato di tripletto<sup>9-11</sup>; esso è composto da un core elettron-ricco dibenzotiofenico, che favorisce il trasporto delle buche, e da due bracci esterni difenilfosforilici (dPPO) elettron-attrattori. Una prima indagine ha coinvolto il piccolo precursore tiofene e i suoi derivati benzo-annulati, benzo[b]tiophene (BBT) e DBT, allo scopo di descrivere la struttura elettronica del tiofene isolato e le modificazioni determinate dalla presenza degli anelli benzo-annulati nel BBT e nel DBT. Ulteriori studi sono stati dedicati al PPT ed ai suoi *building blocks*, DBT e ossido di trifenilfosfina (TPPO), allo scopo di valutare l'impatto di questi ultimi sulla struttura elettronica dell'intero sistema molecolare del PPT; questa indagine era atta inoltre a dimostrare la rottura della coniugazione tra il *core* centrale dibenzotiofenico e le braccia periferiche dPPO, che rende il sistema molecolare del PPT un interessante materiale ambipolare<sup>12</sup>. Un altro obiettivo era quello di fornire una visione più chiara sulla struttura del legame P–O bond, dal momento che la sua descrizione è tuttora controversa<sup>13</sup>. Gli spettri XPS e NEXAFS sono stati simulati alla soglia *K* degli atomi di C e O per mezzo della DFT, usando il metodo del *Transition Potential* (TP), e alle soglie *L<sub>II,III</sub>* degli atomi di S e P mediante la TDDFT con l'approssimazione ZORA.

L'attività di ricerca è stata poi focalizzata su una ftalocianina (Pc) metallica, la titanil ftalocianina (TiOPc), largamente impiegata come semiconduttore organico di tipo *p* nonché come fotoconduttore che trova applicazioni promettenti anche nello sviluppo di OLEDs ad alta prestazione e celle solari<sup>14-17</sup>. L'obiettivo era la determinazione della natura dell'interazione della TiOPc con la superficie di Ag(111) mediante spettri NEXAFS *angle-resolved* alla soglia *K* degli atomi di C, N e O, dal momento che l'orientazione della molecola rispetto alla superficie è tuttora discussa<sup>18-20</sup>. Questo studio è stato condotto in collaborazione con gruppo di Solidi Molecolari dell'Università Philipps di Marburgo.

In seguito, l'attività di ricerca ha avuto per oggetto una serie di sistemi boronici in fase gassosa, l'acido borico [B(OH)<sub>3</sub>], l'acido fenilboronico (PBA), e l'acido 1,4-fenil diboronico (PDBA), importanti *building blocks* dei cosiddetti *Covalent Organic frameworks* (COFs)<sup>21-23</sup>. Questo studio è stato motivato dai risultati ottenuti in un recente lavoro<sup>24</sup>, che accenna all'importanza degli effetti di



correlazione nello stato eccitato del core 1s degli atomi di B di derivati dell'acido boronico. Quindi sono stati calcolati gli spettri NEXAFS alla soglia  $K$  degli atomi di B del  $B(OH)_3$ , PBA, e PDBA attraverso schemi computazionali di complessità crescente, al fine di valutare l'influenza degli effetti di correlazione elettronica; in particolare sono stati impiegati lo schema puramente mono-elettronico dell'approccio DFT-TP<sup>25</sup>, la TDDFT<sup>26</sup>, ed il metodo multiconfigurazionale *self-consistent field* (MCSCF)<sup>27,28</sup>. Gli spettri teorici sono stati confrontati con quelli sperimentali misurati alla *beam line* ALOISA del Sincrotrone Elettra di Trieste.

Ulteriori studi hanno riguardato l'interazione di una serie di networks boronici con la superficie di Au(111), in collaborazione con ricercatori sperimentali afferenti alla *beam line* ALOISA del Sincrotrone Elettra di Trieste. La modellizzazione dell'interfaccia molecola/superficie è stata facilitata da immagini STM (*Scanning Tunnelling Microscopy*) seguita da un'ottimizzazione preliminare della superficie e delle molecole adsorbite attraverso una metodologica *periodic slab* mediante il formalismo DFT<sup>29</sup>. Quindi, opportuni clusters di dimensioni finite sono stati ritagliati per il calcolo degli spettri NEXAFS alla soglia  $K$  degli atomi di B, dal momento che il modello di clusters di dimensioni finite si è dimostrato essere particolarmente adeguato alla rappresentazione di eventi che avvengono sulla superficie o entro il bulk di sistemi solidi, nonché molto efficiente ed accurato per simulare spettri NEXAFS<sup>30-32</sup>. Quindi, la natura dell'interazione adsorbato-superficie è stata discussa in termini di differenze tra il profilo spettrale della molecola in fase gassosa e quello della molecola adsorbita sulla superficie.

La presente Tesi di Dottorato si suddivide in due parti, delle quali la prima (Capitoli 1-2) fornisce una panoramica generale sui concetti teorici, ovvero sugli aspetti principali della spettroscopia NEXAFS, riguardante molecole in fase gassosa ed adsorbite su superfici, e i punti salienti della DFT e della TDDFT, con particolare enfasi sul trattamento degli effetti di rilassamento in seguito alle eccitazioni di core. La seconda parte della Tesi riporta risultati selezionati dell'attività di ricerca ottenuti durante i tre anni di Dottorato che sono stati illustrati in precedenza. Infine verranno illustrati i possibili sviluppi futuri dell'attività di ricerca condotta.

## 1. Core-level spectroscopy

In the last decades, the use of third-generation synchrotrons has boosted the development of a series of spectroscopic methods, among which, core-level spectroscopies have particularly grown in popularity. These spectroscopies are based on the creation of an atomic/molecular excited state characterized by a core-hole; this occurs when the excitation energy of the incident radiation is comparable to the binding energy (BE) of the core levels. The excitation energies fall in the region of *X-Rays*: 100 eV-3 keV (*Soft X-Rays*, in the case of low *Z* atoms), and above 3 keV (*Hard X-Rays*, for heavy atoms, *e.g.*, transition metals). During the absorption of the incident radiation, two different phenomena may occur, namely, the excitation of the core electron to a virtual bound state or to the continuum, *i.e.*, above the ionization threshold. XAS (*X-Ray Absorption Spectroscopy*) and XPS (*X-Ray Photoelectron Spectroscopy*) provide a complementary information on the electronic structure of the investigated system: XAS characterizes the virtual bound states, while XPS provides information on the occupied molecular states, in particular on the BE (*i.e.*, ionization potential) of core electrons.

### 1.1 X-Ray Absorption Spectroscopy (XAS)

XAS (*X-Ray Absorption Spectroscopy*) is a widely used spectroscopic technique, particularly suitable to study the local geometric and/or electronic structure of matter (in gas-phase, solution or condensed). It was first applied to gas phase molecules during the 1930's<sup>33,34</sup>, after the unveiling of the shell structure of atoms. X-ray absorption investigations have known a great development with the introduction of the synchrotron light (electron synchrotron sources, with  $E < 1$  GeV in the 1960s and early 1970s, and electron storage rings, with  $E > 1$  GeV in the 1970s<sup>35</sup>).

XAS is focused on the behavior of the absorption coefficient  $\mu$  of the sample as a function of the incident photon energy. The absorption coefficient of a particular inner shell corresponds to the cross section for a photoexcitation of an electron from that shell. Due to the direct connection between the definition of  $\mu$  and the concept of cross section, the following treatment will be firstly focused on the derivation of an expression for the latter, then, the fundamental aspects of the X-ray absorption process will be described.

#### 1.1.1 The X-Ray Absorption Cross Section

The X-ray absorption cross section ( $\sigma_x$ ) of an atom or molecule is defined as the number of excited electrons per unit time divided by the number of incident photons per unit time per unit area; it is therefore expressed in units of an area, namely, in  $\text{cm}^2$  or barn ( $1 \text{ cm}^2 = 10^{24} \text{ barn}$ ). In the case of photon absorption, one can adopt a semi-classical approach, classically representing the

electromagnetic field but describing the atomic molecular states by means of quantum-mechanics. If one considers only the interaction between the photon field and the electronic charge, and employs first order time dependent perturbation theory, the Fermi's Golden Rule for the transition probability per unit time  $P_{if}$  from the initial state  $|i\rangle$  to the final state  $|f\rangle$  driven by a plane electromagnetic wave can be obtained:

$$P_{if} = \frac{2\pi}{\hbar} |\langle f|H^{(1)}|i\rangle|^2 \rho_f(E) \quad (1.1)$$

where  $\rho_f(E)$  is the energy density of final states, and  $H^{(1)}$  is the interaction Hamiltonian. The latter is dominated by a perturbative term which describes the interaction of the electronic charge with the electromagnetic field:

$$H^{(1)} = -\frac{e}{mc} \mathbf{A} \cdot \mathbf{p} \quad (1.2)$$

where  $e$  and  $m$  are respectively the elemental charge and the mass of the electron, while  $\mathbf{p} = \sum_i \mathbf{p}_i$  is the sum of the linear momentum operators of the electrons.

In the case of *K-shell* excitations,  $P_{if}$  represents the number of electrons excited per unit time from the 1s shell to a final state  $|f\rangle$ , which can be either a bound or a continuum state. The inner shell excitation is driven by a plane electromagnetic wave with electric field vector  $\mathbf{E}$  and vector potential  $\mathbf{A}$ . The latter can be written in the following way:

$$\mathbf{A} = \hat{\mathbf{e}} A_0 \cos(\mathbf{k} \cdot \mathbf{x} - \omega t) = \hat{\mathbf{e}} \frac{A_0}{2} (e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} + e^{-i(\mathbf{k} \cdot \mathbf{x} - \omega t)}) \quad (1.3)$$

where  $\mathbf{k}$  and  $\omega$  are the wave vector and the frequency of the plane electromagnetic wave, respectively, while  $\hat{\mathbf{e}}$  is the unit electric field vector.

By substituting equations (1.2) and (1.3) into Eq. (1.1) and considering that only the time dependent term  $e^{-i\omega t}$  in Eq. (1.3) is responsible for transitions that absorb energy, one can obtain an expression for  $P_{if}$ :

$$P_{if} = \frac{\pi e^2}{2\hbar m^2 c^2} A_0^2 |\langle f|e^{i\mathbf{k} \cdot \mathbf{x}} \hat{\mathbf{e}} \cdot \mathbf{p}|i\rangle|^2 \rho_f(E) \quad (1.4)$$

The dipole approximation (*i.e.*, first order expansion of the exponential) can be efficiently applied to Eq. (1.4); this leads to the following expression for the X-ray absorption cross section:

$$\sigma_x = \frac{4\pi^2 \hbar^2 e^2}{m^2} \frac{1}{\hbar c \hbar \omega} |\langle f|\hat{\mathbf{e}} \cdot \mathbf{p}|i\rangle|^2 \rho_f(E) \quad (1.5)$$

where  $\mathbf{p}$  is the total linear momentum of the electrons. The expression for  $\rho_f(E)$  relies on the normalization of the wave functions; in particular, bound states have unit normalization, whereas continuum states are normalized to a Dirac delta function in the kinetic energy of the photoelectron<sup>36</sup>. The calculation of the spectral resonance features according to Eq. (1.5) requires the wave function of the molecule in its ground state, the wave functions of the bound and/or quasi-bound (continuum) final states, and the energy differences between the ground state and the various resonant excited states. Transition intensities are determined by the wave functions through the dipole matrix element, while the corresponding energies are given by the energy differences.

### 1.1.2 Dipole selection rules

Within the electric dipole approximation, the probability of transition between two electronic states of a system,  $\Psi_i$  (initial state) and  $\Psi_f$  (final state), induced by an oscillating electric field, is proportional to the square modulus of the transition moment,  $M_{if}$ , between the initial and final state:

$$M_{if} = \int \Psi_f^* \boldsymbol{\mu} \Psi_i d\tau \quad (1.6)$$

where  $\boldsymbol{\mu}$  is the electric dipole moment which is defined as:

$$\boldsymbol{\mu} = \sum_i q_i \mathbf{r}_i \quad (1.7)$$

where  $q_i$  and  $\mathbf{r}_i$  are the charge and the  $i$ -th particle position vector, respectively.

From group-theory, a transition is allowed if the direct product between the irreducible representations of the initial state, the final state and the  $\boldsymbol{\mu}$  component contains the total-symmetric representation of the point group of the molecule:

$$\Gamma_f \otimes \Gamma_{\boldsymbol{\mu}} \otimes \Gamma_i \supset \Gamma_{TotSymm} \quad (1.8)$$

Moreover, when spin-orbit interaction can be neglected, the *spin selection rule* requires the conservation of the total electronic spin (*i.e.*,  $\Delta S = 0$ ). Therefore, in the case of *closed-shell* systems, as those investigated in this thesis work, only singlet  $\rightarrow$  singlet transitions are allowed.

As in NEXAFS spectroscopy the electronic transitions start from core orbitals strongly localized on the atomic sites, the dipole selection rule can be expressed as  $\Delta l = \pm 1$ , *i.e.*, transitions are allowed only if the difference between the angular momenta relative to the initial and final state changes by only one unit. The oscillator strengths are directly connected with the atomic site dipole-allowed component of the virtual orbitals. Hence, *K-edge* spectra can be essentially read in terms of  $s \rightarrow p$  transitions, while *L<sub>II,III</sub>-edge* spectra in terms of  $p \rightarrow d$  and  $p \rightarrow s$  transitions (the latter are in most cases 1 or 2 orders of magnitude lesser than the former, especially in the case of transition

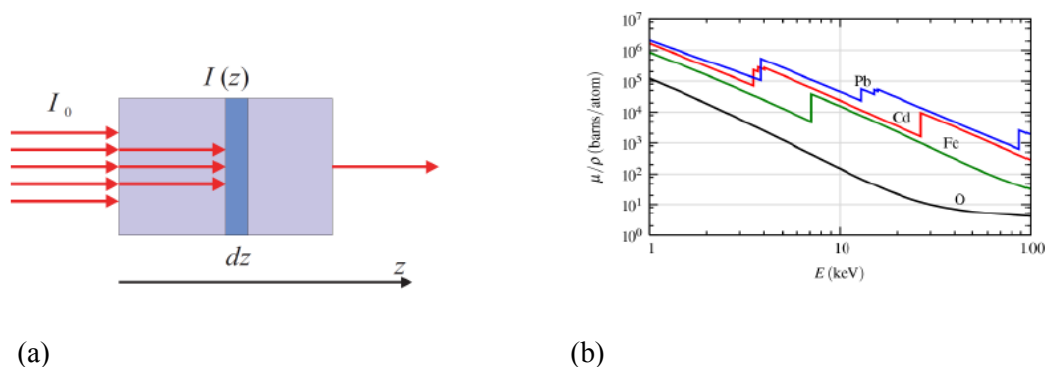
metal compounds); this means that the intensity reflects the amount of the  $p$  and  $s,d$  character of the excited atom in the unoccupied electronic states, respectively.

### 1.1.3 The X-ray Absorption Process

In XAS experiments, the wavelength of the incident X-ray photon is tuned, and the corresponding absorption is measured as a function of the photon energy. The X-ray absorption process is governed by the *Lambeer-Beer law*. When a photon beam of intensity  $I$  passes through an absorbing sample, the infinitesimal intensity loss  $\frac{dI}{I}$  following the absorption process is equal to:

$$\frac{dI}{I(z)} = -\mu dz \quad (1.9)$$

where  $dz$  is the infinitesimal thickness at a depth  $z$  from the surface of the sample (see Figure 1.1).



**Figure 1.1** – (a) Description of the absorption process. (b) Absorption coefficients of Pb, Cd and Fe elements as a function of the incident photon energy.

If  $I(z=0)=I_0$  (with  $I_0$  the incident beam intensity at  $z = 0$ ), the solution of the differential Eq. (1.9) is:

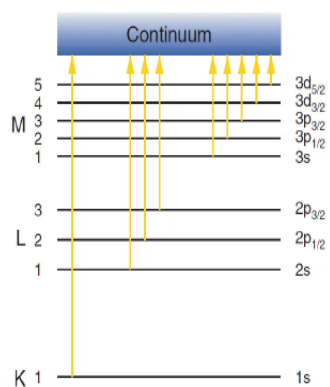
$$I = I_0 \cdot e^{-\mu \cdot x} \quad (1.10)$$

where  $I_0$  is the intensity of the incident light beam,  $I$  is the transmitted intensity, and  $x$  is the thickness of the sample.

Eq. (1.10) represents the fundamental relation of XAS spectroscopy. The absorption coefficient  $\mu$  can be experimentally determined as the ratio of the beam intensities with and without the samples. In non-resonant regions, it is inversely proportional to the energy of the incident photon; this implies that photons with higher energies are more penetrating except at the absorption edges. The latter correspond to the energy of the lowest energy state reached by the core excitations. Approaching to a certain absorption edge, the energy of the photon is not enough to photoemit electrons to the continuum, but is able to promote them to higher empty states: the incident X-ray

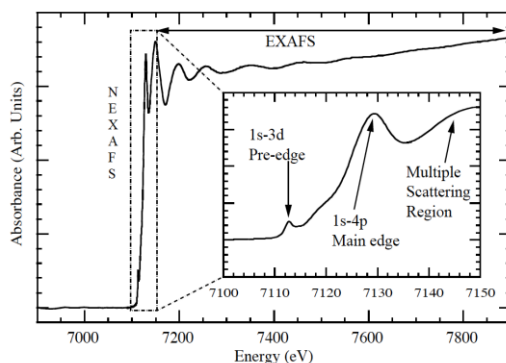
photon energy is scanned through the BE of the core shell, determining an increase in the absorption cross section. It is therefore evident that measuring the absorption cross section in the vicinity of an absorption edge provides direct information on the empty states of the system. At these energies, the material displays a strong X-ray absorption, referred to as *atomic absorption edge*; above the latter, the absorption probability monotonically decreases. The energy at which the electron is ejected into the continuum defines the *ionization threshold* or *ionization potential* (IP).

XAS is involved in the study of  $\mu$  around the absorption edge. Each edge represents a different initial state  $|i\rangle$ , and is conventionally labeled according to the principal quantum number  $n$  of the excited electron (K,L,M,... for  $n = 1, 2, 3...$ ). For instance, K, L<sub>I</sub>, L<sub>II</sub>, L<sub>III</sub>, etc. correspond to the creation of core-holes in the  $1s(^2S_{1/2})$ ,  $2s(^2S_{1/2})$ ,  $2p(^2P_{1/2})$  and  $2p(^2P_{3/2})$ , etc. atomic sub-shells. The Roman numbers stand for the orbital angular momentum,  $l$ , and the total angular momentum,  $J$ , *i.e.*,  $s_{1/2}, p_{1/2}, p_{3/2}, \dots \rightarrow I, II, III, \dots$  (see Figure 1.2).



**Figure 1.2** – Nomenclature of the absorption edges.

The X-ray absorption spectrum is commonly divided into regions (see Figure 1.3), according to the energy of the incident photon, and following on the distinct interpretation of the spectral features and on the kind of the provided information: they are respectively addressed to as NEXAFS (*Near-Edge X-ray Absorption Fine Structure*) and EXAFS (*Extended X-ray Absorption Fine Structure*). To a first approximation, the boundary between these two regions corresponds to the wavelength that equals the distance between the absorbing atom and its first neighbors<sup>37</sup>. NEXAFS is one of the few non-destructive structural probes for non-crystalline and highly disordered materials, such as solutions and complex systems (*e.g.*, polymers). However, the interpretation of NEXAFS spectra is considerably harder than that of the EXAFS ones; therefore, accurate calculations are fundamental for a reliable interpretation of the measured spectral features.



**Figure 1.3** – Representation of NEXAFS and EXAFS energy regions.

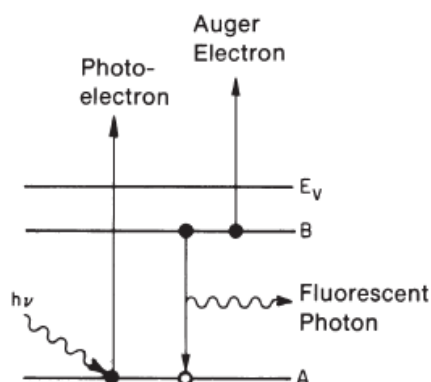
The NEXAFS fine structure observed near the absorption edges was first known as "Kossel structure" from the theory of Kossel<sup>38</sup>. The NEXAFS spectrum is further subdivided into three energy regions: the lower-energy *pre-edge* region, where the incident photon energy is lower than the ionization potential of the core electrons ( $E_0$ ), the region around  $E_0$  (*main-edge*), and the *multiple-scattering* region, in which the incident photon energy is larger than  $E_0$ , and the observed features are essentially explained by means of multiple-scattering resonances. In this thesis work, the analysis of NEXAFS spectra will be exclusively focused on the *pre-edge* region, since the computational protocol adopted is able to accurately characterize only transitions toward virtual bound states.

In 1920, during an investigation on solids, Siegbahn discovered the influence of the chemical environment on the X-ray absorption spectra of bonded atoms<sup>39,40</sup>. Indeed, NEXAFS spectroscopy represents a well-suited tool for obtaining atom-specific and symmetry-selective electronic structure information thanks to the highly localized nature of the core-hole. Moreover, the energy shift of core MOs substantially depends only on the absolute strength of the field produced in the vicinity of the nucleus by the directly bonded neighboring atoms. NEXAFS spectroscopy is also able to provide a useful "mapping" on virtual MOs, according to the atomic contribution on different sites. Therefore, the position and intensity of NEXAFS spectral features are good probes to determine the nature of the virtual states in terms of atomic orbitals (AOs) centered on the atom carrying the core-hole. The near-edge absorption structures can be theoretically analyzed in terms of discrete below edge transitions of the core electron toward virtual orbitals with well-defined excitation energies and intensities, and of continuum transitions which give rise to the photoionization cross section profiles above the ionization threshold. The below edge excitations can be treated by means of the *Linear Combination of Atomic Orbitals* (LCAO) approach in terms of one-electron transitions which usually represent the main contribution to the discrete structures of NEXAFS spectra.

The energy region that starts from the upper limit of NEXAFS and spreads for hundreds of eV past the ionization edge is referred to as EXAFS, and the signal is due to single scattering of the ejected photoelectron; it is usually exploited for the determination of interatomic distances, coordination numbers and species of neighbors surrounding the absorbing atom. EXAFS was firstly denoted as "Kronig structure", in honor of the homonym scientist<sup>41</sup>. In the 1980s, the Kronig near edge structure was mostly referred to as the *X-ray Absorption Near Edge Structure* (XANES)<sup>37</sup> or NEXAFS<sup>42</sup>. The acronyms XANES and NEXAFS are roughly equivalent, even if, over the years, "XANES" has been more often employed for solids and inorganic complexes, while "NEXAFS" has been mostly used for molecules absorbed on surfaces. Currently, "NEXAFS" is by far the most commonly used terminology, and the term "XANES" is employed only to denote the lower-energy NEXAFS region (up to 10 eV above the absorption edge).

### 1.1.4 The fate and nature of core-excited states

In NEXAFS experiments, the sample is irradiated with monochromatic X-rays. The energy of the X-rays is changed in a range of few eV around the ionization threshold of one of the electron shells of the investigated element; in this energy range, transitions toward unoccupied virtual orbitals form a core-hole and a "particle" state. The lifetime of the core-hole is very short: the decay of the excited atomic state generally occurs within  $10^{-15}$  s through *recombination* either radiatively or non-radiatively, involving two competitive processes, namely, *fluorescence* and the *Auger Effect*, respectively (see Figure 1.4).



**Figure 1.4** – Schematic diagram of a photon absorption process resulting in a photoelectron and a core hole. The hole is filled by an electron from a higher shell, either radiatively by emission of a fluorescent photon, or non-radiatively by emission of an Auger electron (adapted from [1]).

When a higher-energy core level electron fills the deeper core-hole and ejects an X-ray photon of well-defined energy, characteristic of the atom, fluorescence occurs; this represents the predominant process for high-Z atoms. The Auger effect instead involves the dropping of an



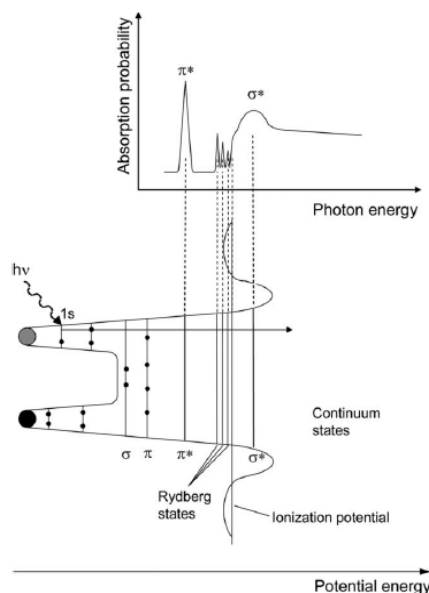
electron from an external shell, and the emission of a second electron into the continuum; the energy due to the electron decay is transferred to an *Auger* electron. The Auger effect is faster (by more than two orders of magnitude) than fluorescence in the case of the *K-shell* excitation of low *Z* atoms and the *L-shell* excitation of all atoms with  $Z < 90$ . Auger electrons are conventionally labelled through three letters, identifying the shell with the core hole, the energetic level of the electron that fills the hole, and the emission shell of the Auger electron, respectively. For instance, the *KLL* nomenclature denotes an Auger electron emitted from the  $n=2$  shell when an electron from the same shell fills a core hole in the  $n=1$  shell. For most samples, the transmitted intensity is difficult to be detected; therefore, the detection of the Auger electrons is preferred. Indeed, electron detection is more surface sensitive with respect to the fluorescence detection due to the short mean free path of electrons in materials.

NEXAFS spectra can be measured in three modes, namely, in *total*, *partial* or *Auger electron yield*. In the total yield mode, all electrons emitted from the sample (*i.e.*, the photoelectrons and the Auger electrons) are revealed by a multi-channel plate detector, while, in the partial yield mode, only electrons of kinetic energy larger than a defined energy threshold are detected; within the latter, a retarding voltage is set in front of the XAS detector to retard the low-kinetic energy electrons. Finally, in the Auger yield mode, a dispersive-type electron energy spectrometer is employed by choosing an appropriate energy window defined by the kinetic energy of the Auger transitions of the element of interest. Due to the different mean free path between the electrons emitted from the surface and fluorescent photons, the most surface-sensitive techniques are those based on the recording of NEXAFS spectra in partial and Auger electron yield modes.

The *K-shell* NEXAFS spectra of atoms and molecules show a variety of resonances arising from electronic transitions of a  $1s$  core electron to delocalized final states near the vacuum level; the interpretation of these spectra is favored by the particularly simple symmetry and localized nature of the  $1s$  initial state. Moreover, core levels are energetically separated from each other, so that one can observe a threshold at a time; the same is not valid in the case of valence absorption spectra, where the electrons are excited from several valence MOs lying close together, resulting in the overlap and interference of lines with one another. As the interaction of the electron hole in the atomic core with the excited electron is relatively small, NEXAFS spectra can be interpreted within the one-electron model, *i.e.*, by treating the core electron as “active” and the  $N-1$  remaining electrons as “passive” (*i.e.*, not involved in the excitation process). As pointed out by Dehmer et al.<sup>43</sup>, the analysis of valence shell spectral features is instead hampered by the delocalized nature of

the valence-hole state, leading to electron correlation and screening effects as well as to strong particle-hole interactions.

A typical *K-shell* NEXAFS spectrum of a diatomic unsaturated molecule (e.g., CO) is represented in the upper panel of Figure 1.5.



**Figure 1.5** – Schematic potential and corresponding *K-shell* NEXAFS spectrum of a diatomic unsaturated molecule. Upper panel: spectral features arising from the core-electron excitation toward empty states. Lower panel: occupied and virtual bound states (adapted from [44]).

Virtual MOs can be divided into two classes, namely, MOs deriving from the valence electron shells of the atoms, which usually have antibonding character, and MOs lying outside the valence electron shells (*Rydberg* orbitals, *i.e.*, orbitals having non-bonding character, with principal quantum number greater than that of any occupied orbital of the ground state). Moreover, unoccupied MOs can be labeled as  $\sigma$  or  $\pi$  according to their spatial displacement and orientation, along and perpendicularly to the bond between two atoms, respectively. In the case of *K-shell* NEXAFS spectra, the observed resonances are associated to dipole-allowed transitions from an initial  $1s$  state, which always has  $\sigma$  symmetry, to a final state of  $\sigma$  or  $\pi$  symmetry. In *K-shell* NEXAFS spectra of low- $Z$  molecules with  $\pi$  bonding, the lowest-energy structure is associated to  $\pi^*$  resonances, and its energy position falls below the ionization threshold, as a result of the increased electron-hole Coulomb attraction. The core-hole is responsible of an upward (*i.e.*, smaller BE) relaxation shift of the  $1s$  orbital as well as of a downward shift of the outer excited orbitals.  $\pi^*$  resonances usually display very sharp peaks due to the localized nature of these states and to the fact that the lifetime of the final state is determined by the de-excitation of the excited electron.

Between  $\pi^*$  resonances and the ionization threshold, several sharp and weak resonances can be visible; these transitions can be either associated to pure Rydberg MOs or to a mixture of Rydberg and H-derived antibonding MOs, in the presence of bonds to H atoms in the molecule. Two observations on the intensities of the implicated transitions are in order. In the energy region close to the ionization threshold, Rydberg features are usually weaker; this can be rationalized by considering the decreased size of the 1s orbital, which leads to a smaller spatial overlap with the Rydberg orbital. On the other hand, the more compact nature of the valence orbitals increases the intensity of transitions toward mixed valence H-derived-Rydberg MOs.

Finally, in the higher-energy region of the spectrum,  $\sigma^*$  resonances can be found; for most low-Z molecules, these usually occur above the ionization threshold, but, when the sum of Z of the two bonded atoms is higher than 15 (e.g., O<sub>2</sub>), they fall below it<sup>45</sup>. Broader peaks are expected for these resonances due to the shorter lifetime of the final state determined by the increased decay probability of the electron to continuum states. Moreover, asymmetric broadenings deriving from the vibrational motion of atoms are observed for these resonances.

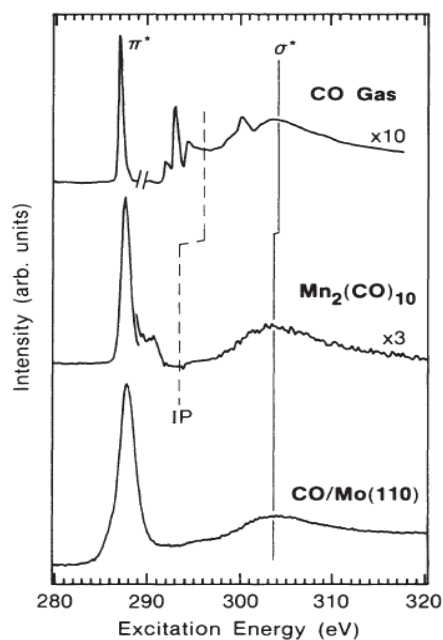
### **1.1.5 NEXAFS spectroscopy of adsorbed molecules**

Another precious information provided by NEXAFS spectroscopy regards the extent of the modifications induced on the electronic and molecular structure of adsorbed molecules by the interaction with a surface. The employment of NEXAFS spectroscopy to determine the molecular orientation is particularly well-established, especially by means of the polarized XAS. In this respect, the comparison between the NEXAFS spectral profiles of the free and adsorbed molecule is particularly useful to determine the nature of the molecule–substrate interaction. Although NEXAFS profiles are still dominated by intramolecular excitations, the influence of the substrate on both positions and intensities of the transitions allows an interpretation of the adsorption process in terms of electronic structure modifications.

The adsorbates can bind to the surface either through a physical (*physisorption*) or a chemical interaction (*chemisorption*). Physisorption occurs when the surface-molecule interaction is weak (*i.e.*, mediated via long-range van der Waals forces), while chemisorption is a strong and site-specific process. In physisorption, the overlap between adsorbate and substrate orbitals is small; therefore, only minor changes in the resonance positions relative to the gas phase spectrum are expected. On the other hand, in chemisorption, the overlap between adsorbate and substrate leads to a re-hybridization of the valence levels with consequent modifications of the spectral profile of the gas phase molecule. The first measurement of a NEXAFS spectrum of a chemisorbed molecule was

that of CO on the Ni(100) surface, and was performed at the Stanford Synchrotron Radiation Laboratory in 1980; the related results were published in conjunction with those relative to NO on the same surface<sup>46</sup>. This publication was focused on the polarization dependence of the near edge resonances, which was used to determine the molecular orientation on the surface; a more detailed study also involved N<sub>2</sub> on Ni(100) and reported the derivation of the equations for the angular dependence of the resonances<sup>42</sup>.

The spectrum of the free molecule is usually not dramatically affected by the extra-molecular interaction between the molecule and the surface upon chemisorption. Extra-molecular effects can be essentially investigated in two steps, namely, by employing condensed or physisorbed molecules, where van-der-Waals type inter-molecular or molecule-surface bonds are very weak with respect to covalent intra-molecular bonds, and chemisorbed molecules or organometallic compounds, where the molecule-metal bonds can be much weaker or comparable to intra-molecular bond strengths. An example of extra-molecular interaction analysis is shown in Figure 1.6 for the carbon monoxide (CO) molecule in different chemical environments. In particular, the C *K-shell* NEXAFS spectrum of gas phase CO (upper panel) is shown together with those of CO in Mn<sub>2</sub>(CO)<sub>10</sub> (middle panel) and strongly bonded via the C atom to the Mo(100) surface, with the C-O axis along the normal to the surface (lower panel).



**Figure 1.6** – C *K-shell* NEXAFS spectra of CO in different chemical environments: gas phase (upper panel), Mn<sub>2</sub>(CO)<sub>10</sub> (middle panel) and strongly chemisorbed on the Mo(110) surface (lower panel) (adapted from [1]).

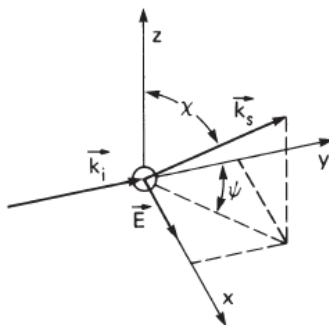
Before analyzing the evolution of the spectral shape in passing from the gas phase to the chemically bonded CO, some important aspects regarding the bonding of CO to metal atoms in carbonyls and on surfaces have to be considered. The bonding is formed by the interaction of the  $5\sigma$  HOMO and  $2\pi^*$  LUMO orbitals with metal orbitals (*Blyholder* model<sup>47,48</sup>). It consists of a donor bond formed by the C  $5\sigma$  lone pair orbital; this bond is stabilized by the back-bonding of metal  $d\pi$ - $p\pi$  orbitals with the  $2\pi^*$  antibonding orbitals of the molecule. This implies that the nature of the metal-molecule back-bonding interaction sensitively affects the  $\pi^*$  resonance. Furthermore, the width of the  $\pi^*$  resonance is associated to the strength of the back-bond as well as to the nature of the metal states. In particular, if the bond strength increases, the metal  $d\pi$ - $p\pi$  contribution to the  $2\pi^*$  orbital will also increase, while the  $\pi^*$  resonance intensity will decrease.

From a general overview of Figure 1.6, all three spectra display a similar shape, being all characterized by a strong  $\pi^*$  resonance around 287.5 eV and a  $\sigma^*$  shape resonance around 304 eV; the energy position of the  $\pi^*$  resonance is therefore not substantially changed by the extra-molecular interactions. This can be rationalized by the fact that bonding shifts of the 1s level associated with bonding to a transition metal are small<sup>49-51</sup>. In the CO/Mo(100) spectrum, the  $\pi^*$  resonance appears broadened; this is due to the fact that the excited electron can decay into the conduction band, thus reducing the lifetime of the excited state. Moreover, in that spectrum, one can notice, on the low energy side of the  $\pi^*$  resonance, the presence of a shoulder deriving from transitions to metal conduction band states. In the case of chemically bonded CO, Rydberg resonances are not visible. Another interesting aspect is the lowering of the IP value in the chemisorbed CO case; this can be ascribed to extra-molecular relaxation effects, *i.e.*, the final hole state is well screened, either by metal conduction electrons or by the excited 1s electron in the  $\pi^*$  state.

NEXAFS spectroscopy can be used to determine the orientation of molecules with respect to a surface by means of angular resolved spectra<sup>44</sup>. The strong directionality of the covalent bonds between low-Z atoms in combination with the polarized nature of the synchrotron radiation leads to a strongly polarization-dependence of  $\pi^*$  and  $\sigma^*$  resonances (*dichroism*) in *K-shell* NEXAFS spectra, even in the case of chemisorbed molecules, although the overall symmetry of the adsorbate complex is generally different from that of the gas phase molecule.

Before discussing the polarization-dependence of resonances, it is useful to point out that light can be linearly polarized in two different modes, denoted as *s* and *p*, respectively. The *s*-polarization (from the German word *senkrecht*, which means “perpendicular”) occurs when the electric field

vector is perpendicular to the incidence plane (the plane that contains the propagation direction of the incident wave and the normal to the interface), whereas the *p*-polarization (from the German word *parallele*, which means “parallel”) is observed when the electric field vector is parallel to the incidence plane (see Figure 1.7).



**Figure 1.7** – Coordinate system for a coherent X-ray scattering process from an atom for linearly polarized X-rays.  $\mathbf{k}_i$  is the incident wave vector along the y-axis,  $\mathbf{k}_s$  is the scattered wave vector. The  $\mathbf{E}$  vector lies along the x-axis. The *s*-polarization is observed for  $\psi = 0^\circ$ , while the *p*-polarization for  $\psi = 90^\circ$  (adapted from [1]).

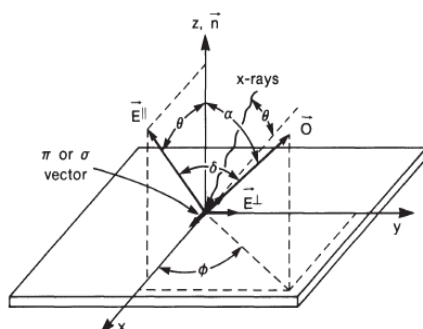
Within a MO approach, one can correlate the polarization dependence of the resonance intensities to the spatial displacement of the final MO (*i.e.*, the direction of maximum orbital amplitude on the excited atom relative to the nuclear framework). In particular, there is a *one-to-one* correlation between the molecular geometry and the spatial orientation of the final MO; this implies that a single orbital can be represented by a vector, and two orthogonal orbitals by a plane. Moreover, it follows that the intensities of resonances associated to final states of different symmetry display a strong and opposite angular dependence. Furthermore, the orientation of the electric field vector of the incident light is thought to be not affected by the surface. The angular dependence is contained in the dipole matrix element  $\langle f | \mathbf{e} \cdot \mathbf{p} | i \rangle$ ; according to Eq. (1.5) and assuming that the X-rays are linearly polarized in the direction of the unit electric field vector  $\mathbf{e}$ , one obtains for the transition intensity associated to the dipole matrix element:

$$I \propto |\mathbf{e} \cdot \langle f | \mathbf{p} | i \rangle|^2 \quad (1.11)$$

where  $|i\rangle$  is the 1s initial state, and  $|f\rangle$  is the MO final state of the transition, while  $\mathbf{p}$  stands for the momentum operator.

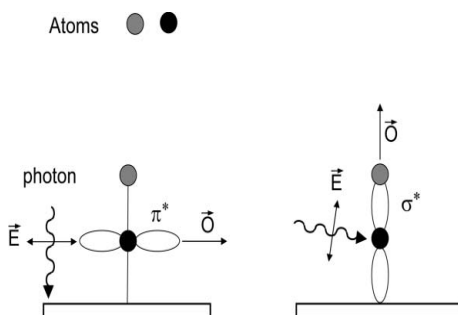
In the case of *K-shell* excitations, the vector matrix element  $\langle f | \mathbf{p} | i \rangle$  points toward the direction of the  $\mathbf{p}$  component in the final state orbital on the excited atom; therefore, the polarization dependence of the total matrix element can be expressed as a function of the angle  $\delta$  between the electric field vector  $\mathbf{E}$  and the direction  $\mathbf{0}$  of largest amplitude of the final state orbital (see Figure 1.8):

$$I_v \propto |\mathbf{e} \cdot \langle f | \mathbf{p} | i \rangle|^2 \propto |\mathbf{e} \cdot \mathbf{0}|^2 \propto \cos^2 \delta. \quad (1.12)$$



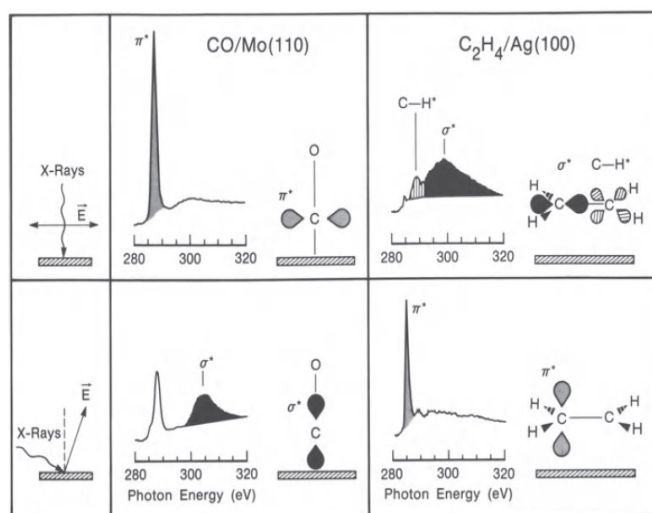
**Figure 1.8** – The orientation of the orbital, *i.e.*, of the vector **O**, is characterized by a polar angle  $\alpha$  and an azimuthal angle  $\phi$ . The X-rays are incident in the (x, z) orbit plane ( $\Theta$  is the incident angle) (adapted from [1]).

In the case of *K-shell* spectra, the dipole selection rule can be therefore generalized as follows: if the electric field vector of the linearly polarized X-ray beam points along the direction of maximum amplitude of the final MO, the intensity of the resonance is largest, while it vanishes when the electric field vector is perpendicular to the final MO; if the electric field vector lies in the nodal plane of the final MO, the transition is forbidden. Hence, in the case of a  $\pi$ -bonded diatomic molecule adsorbed with its molecular axis normal to the surface, the  $\pi^*$  resonance is maximized at normal incidence (*i.e.*, parallel to the bond axis), while the  $\sigma^*$  resonance is greatest at grazing incidence (see Figure 1.9).



**Figure 1.9** – Schematic representation for the angular dependency of  $\pi^*$  and  $\sigma^*$  resonances in the NEXAFS spectrum for a  $\pi$ -bonded diatomic molecule adsorbed with its molecular axis normal to the surface (adapted from [44]).

The dependence of the resonance intensities on the orientation of the electric field vector relative to the molecular axes is displayed in Figure 1.10 for the C *K-shell* NEXAFS spectra of CO on Mo(110) and of ethylene ( $C_2H_4$ ) on Ag(100). The spectra of the two chemisorbed systems are recorded at normal X-ray incidence (*i.e.*, when the electric field vector lies in the surface plane, upper panels) and grazing X-ray incidence angle (*i.e.*, when the electric field vector form a  $20^\circ$  angle from the surface normal, lower panels).



**Figure 1.10** – Polarization dependence of resonances in C *K-shell* NEXAFS spectra. Left side: CO on Mo(110), right side: C<sub>2</sub>H<sub>4</sub> on Ag(100) (adapted from [1]).

In the following treatment, the electric field vector is used as a pointer, and the X-rays are assumed to be perfectly linearly polarized. Starting with CO on Mo(110), when the electric field vector is parallel to the surface, only one peak around 287 eV, corresponding to the  $\pi^*$  resonance, is found, while, for the electric field vector nearly perpendicular to the surface, this peak is significantly less intense, and a higher-energy peak, associated to the  $\sigma^*$  resonance, is visible around 304 eV; these considerations indicate that the molecule stands up on the surface. In analyzing the recorded spectra for C<sub>2</sub>H<sub>4</sub> on Ag(100), one has to keep in mind that, unlike CO, the C<sub>2</sub>H<sub>4</sub> molecule typically bonds to metal atoms and surfaces in a *side-on* geometry, due to its preference to form " $\pi$  bonds" to metal atoms<sup>52</sup>. Similarly to CO, the bonding is made of two contributions, namely, a donor bond between the occupied C-C  $\pi$  orbital of the molecule and a  $\sigma$  orbital of the metal, and an acceptor bond in which the metal  $d\pi-p\pi$  and  $\pi^*$  antibonding orbitals of the molecule are implicated<sup>53,54</sup>. The most interesting aspect regarding the two spectra relative to C<sub>2</sub>H<sub>4</sub> on Ag(100) concerns the opposite angular dependence of the  $\pi^*$  and  $\sigma^*$  resonances (both associated to the C=C double bond of the molecule) to that of CO. Indeed, since the  $\pi^*$  resonance is found for the electric field vector nearly perpendicular to the surface, and the  $\sigma^*$  resonance for the electric field vector parallel to the surface, one can conclude that the molecule is lying down on the surface, *i.e.*, with the C-C axis parallel to it. In the spectrum recorded for the electric field vector parallel to the surface, a resonance associated to  $\sigma^*(\text{C-H})$  states is visible around 288.5 eV. Since the intensity of this peak is higher for this polarization, this implies that the C-H bonds lie in a plane parallel to the surface (or close to that plane); therefore, the molecule orients itself on the surface.



## 1.2 X-Ray Photoelectron Spectroscopy (XPS)

Photoelectron spectroscopy is a well-suited method to investigate the electronic structure of molecules, solids and surfaces, and has been employed in several fields, *e.g.*, surface chemistry and material science, providing important contributions to the understanding of fundamental principles in solid state physics. It was developed by Siegbahn, who originally introduced it as *Electron Spectroscopy for Chemical Analysis* (ESCA) to point out its importance in surface chemistry<sup>55</sup>.

This technique is based on the *photoelectric effect*, which was discovered by Einstein in 1905<sup>56</sup>; this process involves the emission of electrons from a sample following the interaction with photons of X-ray energy (in the order of the keV). The initial *N-electron* state evolves toward a final *N-1 electron* bound state (ion) and a free electron in the electronic *continuum*. The one-electron approximation is adequate to describe the photoemission model; however, it is not able to take into account any interaction between the hole, formed upon the photoemission process, and the surrounding electrons. The main difference between the final state and the initial one lies, to a first approximation, in the presence of the hole; this corresponds to employ the *N* one-electron wave functions (orbitals) solutions of the one-particle Schrödinger equation to build the Slater determinants of both initial and final states (characterized by *N* and *N-1* electrons, respectively). The BE of the core electron is given by the difference between the total energies associated to the two Slater determinants. According to Koopmans' theorem<sup>57</sup>, the BE of the core electron corresponds, in modulus, to the energy of the core orbital. However, the effective photoelectron BE does not correspond to that predicted by Koopmans' theorem, since the energy of the final state is reduced by relaxation phenomena induced by the creation of the core-hole; this aspect will be further discussed in the Theoretical methods Section (see Paragraph 2.5.7).

The quantity that is actually measured in XPS experiments is the kinetic energy (*KE*) of the emitted photoelectron as a result of the core-hole formation. During the photoemission process, the sample is irradiated with a monochromatic electromagnetic radiation of frequency  $\nu$ ; the latter is kept fixed, and must be higher than the BE of the core electron. This implies that the following energetic balance must be satisfied:

$$KE = h\nu - BE \quad (1.13)$$

For samples in gas phase, the BE corresponds to the IP of the emitted core electron. The latter is defined as the minimum energy to excite a core-electron to the continuum of states above the vacuum level, and can be calculated as the energy difference between the *N-1* electron and the *N*-electron systems, as follows:

$$IP = E^{N-1} - E_0^N \quad (1.14)$$

For solid phase samples, one must consider the *work function*,  $\phi$ , which represents the minimum energy to extract an electron from a surface to the vacuum level with zero kinetic energy. Therefore, the corresponding energetic balance can be written as:

$$KE = h\nu - \phi \quad (1.15)$$

The core-excited state relaxes upon ionization, and the vacancy is filled by an electron from a higher level which can lead to either X-ray fluorescence or Auger emission.

Non-equivalent atoms of the same element in a molecule give rise to distinct peaks at different BEs. From core BEs, it is possible to obtain physical and chemical information (*e.g.*, the oxidation state) of the investigated atomic species. The BE chemical shifts are influenced by the chemical and geometric environment of the ionized atomic site, in terms of electronic density (initial state effects) and electronic relaxation (final state effects). Generally, a different electronic density distribution on non-equivalent atoms causes a shift on the corresponding IP. In particular, a decrease of the electronic density leads to an increase of the IP (de-shielding effect), whereas an increase of the electronic density corresponds to a decrease of the IP (shielding effect); these trends can be rationalized by the different effective nuclear charge ( $Z_{eff}$ ) felt by the core-electrons according to the different charge densities. For instance, the presence of an electronegative atom (*e.g.*, O) increases the IP of a less electronegative atom (*e.g.*, C) bonded to it: as a result of the electronic attraction, the electrostatic shielding at the C nucleus decreases with a consequent increase of the corresponding IP.

Furthermore, XPS can be exploited as surface-sensitive technique in the study of the outermost atomic layers of surfaces. The great surface-sensitivity derives from the short mean free path of the low-energy electrons in the materials; indeed, for an energy interval in the range from some eV to keV, the mean free path is reduced to few tens of Å. The energy of the photons typically employed in XPS experiments is up to 1500 eV; the penetrating power of these photons in solid surfaces is of the order of 1-10  $\mu\text{m}$ .<sup>58</sup> The interaction of photoemitted electrons with the solid surface allows for a sensitivity of only some tens of Å.<sup>59</sup> The propagated photoelectrons in the solid can be elastically or in-elastically scattered: in the first case, the electrons escape the solid without losing their kinetic energy, while in the second one, electrons lose part of it in coming out from the solid. A peculiar aspect to consider in XPS experiments on surfaces is that core levels often have a BE that differs between bulk and surface atoms due to the changed potential at the surface<sup>58</sup>; this usually determines the detection of two distinct components in the substrate photoemission peaks.

## 2. Theoretical methods

### 2.1 Hartree-Fock method

Since the birth of quantum mechanics, the search for approximate solutions to the electronic Schrödinger equation has been of great interest for quantum chemists. In the resolution of many-electron problems, the Hartree-Fock (HF) method has played an important role, elucidating the bases of modern quantum chemistry. The central idea of the HF approximation is to reduce the complicated many-electron problem to a one-electron problem in which the repulsion among electron pairs is treated in an average way. Electrons are treated as independent particles; therefore, electronic correlation is not taken into account. To include electron correlation in the multi-electronic wave function, the HF method represents a good starting point for more elaborate many-body theories (*e.g.*, *Configuration Interaction*, *Coupled Cluster* and *Moller-Plesset perturbation theory*). A robust alternative to HF methods used in the case of very large systems is represented by Density Functional Theory (DFT), which treats both exchange and correlation energies, although in an approximate way. Indeed, it is common to perform calculations that are a hybrid of the two methods, such as the popular B3LYP scheme<sup>60</sup>.

Within the HF approach, according to the independent-particle approximation and the Pauli exclusion principle, the ground state of an  $N$ -electron system described by an electronic Hamiltonian  $H$  can be built through a single Slater determinant, *i.e.*, an antisymmetric product of individual electron spin orbitals<sup>61</sup>:

$$|\Psi_0\rangle = |\chi_1\chi_2 \dots \chi_N\rangle. \quad (2.1)$$

According to the variational principle, the best wave function of this functional form is the one which gives the lowest energy:

$$E_0 = \langle\Psi_0|H|\Psi_0\rangle, \quad (2.2)$$

where  $H$  denotes the full electronic Hamiltonian. The problem therefore consists in finding a set of spin orbitals  $\{\chi_a\}$  such that  $|\Psi_0\rangle$  corresponds to the best approximation to the ground state of the  $N$ -electron system. The spin orbitals  $\{\chi_a\}$  can be systematically varied until the minimum of the energy  $E_0$  is reached, with the constraint of preserving the orthonormality condition:

$$\langle\chi_a|\chi_b\rangle = \delta_{ab} \quad (2.3)$$

In this way, it is possible to derive an equation, whose solutions correspond to the best spin orbitals (*i.e.*, the ones that minimize  $E_0$ ). This equation is the HF integro-differential equation, and has the following form:

$$h(1)\chi_a(1) + \sum_{b \neq a} \left[ \int d\mathbf{x}_2 |\chi_b(2)|^2 r_{12}^{-1} \right] \chi_a(1) - \sum_{b \neq a} \left[ \int d\mathbf{x}_2 \chi_b^*(2) \chi_a(2) r_{12}^{-1} \right] \chi_b(1) = \varepsilon_a \chi_a(1) \quad (2.4)$$

where  $h(1) = -\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{r_{1A}}$  is the kinetic energy and potential energy for attraction to the nuclei experienced by electron 1, while  $\varepsilon_a$  is the energy of the spin orbital  $\chi_a$ . It is worth noting that the two terms in Eq. (2.4) involving sums over  $b$  are those that represent the repulsion among electron pairs in single determinant HF theory; in particular, the first one is the *Coulomb* term, while the second one is the *exchange* term, which derives from the antisymmetric nature of the electronic wave function. Without these two terms, Eq. (2.4) is reduced to a one-electron Schrödinger equation for the spin orbital states of a single electron in the field of the nuclei:

$$h(1)\chi_a(1) = \varepsilon_a \chi_a(1). \quad (2.5)$$

The *Coulomb* term represents the total averaged one-electron coulomb potential acting on electron 1 in  $\chi_a$ , arising from the  $N-1$  electrons in the other spin orbitals; it is thus convenient to define a *Coulomb* operator:

$$J_b(1) = \int d\mathbf{x}_2 |\chi_b(2)|^2 r_{12}^{-1}, \quad (2.6)$$

which represents the average local potential at  $\mathbf{x}_1$  arising from an electron in  $\chi_b$ . Therefore, the effect of  $J_b(1)$  when operating on a spin orbital  $\chi_a(1)$  is the following:

$$J_b(1)\chi_a(1) = \left[ \int d\mathbf{x}_2 \chi_b^*(2) r_{12}^{-1} \chi_b(2) \right] \chi_a(1). \quad (2.7)$$

By introducing an exchange operator  $K_b(1)$  acting on a spin orbital  $\chi_a(1)$ , the HF integro-differential equation (Eq. 2.4) can be rewritten as:

$$[h(1) + \sum_{b \neq a} J_b(1) - \sum_{b \neq a} K_b(1)]\chi_a(1) = \varepsilon_a \chi_a(1), \quad (2.8)$$

where the effect of  $K_b(1)$  when operating on a spin orbital  $\chi_a(1)$  is the following:

$$K_b(1)\chi_a(1) = \left[ \int d\mathbf{x}_2 \chi_b^*(2) r_{12}^{-1} \chi_a(2) \right] \chi_b(1). \quad (2.9)$$

The operator in square brackets in Eq. (2.8) is different for every spin orbital  $\chi_a$  on which it operates (due to the restricted summation over  $b \neq a$ ). Now, considering the effect of the Coulomb and exchange operators on the spin orbital  $\chi_a(1)$  (Eqs. 2.7 and 2.9), it follows that:

$$[J_a(1) - K_a(1)]\chi_a(1) = 0. \quad (2.10)$$

It is therefore possible to add this term to Eq. (2.8), eliminating the restriction on the summation, and defining a *Fock operator*  $f$  as:

$$f(1) = h(1) + \sum_b [J_b(1) - K_b(1)]. \quad (2.11)$$

The Fock operator therefore corresponds to the sum of the *core-Hamiltonian operator*  $h(1)$  and the HF potential  $v^{HF}(1)$ , which represents the average potential experienced by electron 1 due to the presence of the other electrons:

$$v^{HF}(1) = \sum_b [J_b(1) - K_b(1)]. \quad (2.12)$$

The HF equations now assume their usual form:

$$f|\chi_a\rangle = \varepsilon_a|\chi_a\rangle, \quad (2.13)$$

that is an eigenvalue equation with the spin orbitals as eigenfunctions and their energy as eigenvalues; the exact solutions correspond to the “exact” HF spin orbitals.

However, since  $v^{HF}(1)$  depends on the spin orbitals (*i.e.*, the Fock operator depends on the Fock orbitals), Eq. (2.13) is nonlinear and must be solved iteratively; this procedure is known as *Self-Consistent Field* (SCF). It consists in making an initial guess at the spin orbitals to calculate the average field seen by each electron, and then solve the eigenvalue equation (Eq. 2.13) by considering a new set of spin orbitals. By using the latter, a new  $v^{HF}$  can be obtained; the procedure is repeated until self-consistency, namely, when  $v^{HF}$  does not change further, and the spin orbitals employed to build the Fock operator correspond to its eigenfunctions.

The solution of Eq. (2.13) provides a set  $\{\chi_k\}$  of orthonormal HF spin orbitals with orbital energies  $\{\varepsilon_k\}$ . The  $N$  spin orbitals with the lowest energies are denoted as *occupied* or *hole* spin orbitals, and are used to build  $|\Psi_0\rangle$ , which, in this way, corresponds to the best approximation to the ground state of the  $N$ -electron system. The remaining members of the set  $\{\chi_k\}$  are called *virtual*, *unoccupied*, or *particle* spin orbitals. Although Eq. (2.13) has an infinite number of solutions and

virtual spin orbitals, it is in practice solved by introducing a finite set of basis functions  $\{\phi_\mu\}$  to expand the spin orbitals, and then solving matrix eigenvalue equations for the expansion coefficients. It is evident that the larger and more complete the set of basis functions  $\{\phi_\mu\}$  is, the greater will be the flexibility in the expansion for the spin orbitals, and the lower will be the expectation value of  $E_0$  (Eq. 2.2). Therefore, as one approaches the HF limit, the obtained spin orbitals approach the “exact” HF spin orbitals.

Finally, it is necessary to point out that the HF wave function usually gives 99% of the exact energy; the difference between the exact nonrelativistic ground state energy of the system and the HF-limit energy corresponds to the *correlation energy*. Moreover, the HF approximation breaks down in cases where degeneracies or near-degeneracies occur between electronic configurations, such as in bond dissociation processes, and along the reaction path for symmetry-forbidden chemical reactions; it therefore represents a valid approximation for most molecules around their equilibrium geometry, but cannot be in general employed as a qualitatively correct model for potential energy surfaces. In the next Paragraphs, methods including electron correlation in the multi-electron wave function, such as the Configuration Interaction (CI), Multi-Configuration Self-Consistent Field (MCSCF), and Complete Active Space Self-Consistent Field (CASSCF) will be outlined.

### 2.1.2 Koopmans’ theorem

Koopmans’ theorem provides a way of calculating approximate ionization potentials and electron affinities by assuming that, upon the removal of an electron from a MO to form a positive ion, the MOs of the ion are the same to those of the neutral molecule (*i.e.*, *frozen orbital approximation*).<sup>57</sup> The main advantage in using this theorem relies on the fact that it is not necessary to calculate the two separate energies of the neutral molecule and of its corresponding ion in order to determine the ionization energy and electron affinity of a molecule.

**Koopmans’ theorem:** Given an  $N$ -electron HF single determinant  $|\Psi_0\rangle$  with occupied and virtual spin orbital energies, then the ionization potential to produce an  $(N-1)$ -electron single determinant  $|\Psi_a^{N-1}\rangle$  with identical spin orbitals, obtained by removing an electron from spin orbital  $\chi_a$ , and the electron affinity to generate an  $(N+1)$ -electron single determinant  $|\Psi_r^{N+1}\rangle$  with identical spin orbitals, obtained by adding an electron to spin orbital  $\chi_r$ , just correspond to the negative of the orbital energy  $\varepsilon_a$  of the spin orbital  $\chi_a$ , and of the orbital energy  $\varepsilon_r$  of the spin orbital  $\chi_r$ .

**Proof:**

Let's first consider the removal of an electron from the  $N$ -electron state  $|^N\Psi_0\rangle = |\chi_1\chi_2 \dots \chi_c \dots \chi_N\rangle$ . The removal of an electron from the spin orbital  $\chi_c$  generates a  $(N-1)$ -electron single determinant state  $|^{N-1}\Psi_c\rangle = |\chi_1\chi_2 \dots \chi_{c-1}\chi_{c+1} \dots \chi_N\rangle$ , where the remaining  $N-1$  spin orbitals in  $|^{N-1}\Psi_c\rangle$  are identical to those in the  $N$ -electron state  $|^N\Psi_0\rangle$ . In second quantization, this is done by annihilating an electron in  $\chi_c$ , so that:

$$|^{N-1}\Psi_c\rangle = a_c|^N\Psi_0\rangle. \quad (2.14)$$

The ionization potential of  $|^N\Psi_0\rangle$  for this process is given by:

$$\text{IP} = {}^{N-1}E_c - {}^NE_0, \quad (2.15)$$

where  ${}^{N-1}E_c$  and  ${}^NE_0$  represent the expectation values of the energy of the two relevant single determinants:

$${}^{N-1}E_c = \langle {}^{N-1}\Psi_c | H | {}^{N-1}\Psi_c \rangle, \quad (2.16)$$

$${}^NE_0 = \langle {}^N\Psi_0 | H | {}^N\Psi_0 \rangle. \quad (2.17)$$

The energy of a single determinant is:

$$E = \sum_i {}^{occ} \langle i | h | i \rangle + \frac{1}{2} \sum_i {}^{occ} \sum_j {}^{occ} \langle ij || ij \rangle, \quad (2.18)$$

where the summations go over all spin orbitals occupied in the determinant. Therefore,

$${}^NE_0 = \sum_a \langle a | h | a \rangle + \frac{1}{2} \sum_a \sum_b \langle ab || ab \rangle, \quad (2.19)$$

where the indexes  $a, b, \dots$  denote the spin orbitals occupied in  $|^N\Psi_0\rangle$ , and:

$${}^{N-1}E_c = {}^NE_0 - \langle c | h | c \rangle - \frac{1}{2} \sum_{a \neq c} \langle ac || ac \rangle - \frac{1}{2} \sum_{b \neq c} \langle cb || cb \rangle = {}^NE_0 - \langle c | h | c \rangle - \sum_{a \neq c} \langle ac || ac \rangle. \quad (2.20)$$

The ionization potential is therefore given by:

$$\text{IP} = {}^NE_0 - \langle c | h | c \rangle - \sum_{a \neq c} \langle ac || ac \rangle - {}^NE_0 = -\langle c | h | c \rangle - \sum_{a \neq c} \langle ac || ac \rangle = -\varepsilon_c. \quad (2.21)$$

Hence, in the single determinant and frozen orbital approximation, occupied spin orbital energies represent the opposite of the energy required to remove an electron from that spin orbital. Since orbital energies of occupied states are negative, ionization potentials are positive.

Let's consider now the process of adding an electron to one of the virtual spin orbitals  $\chi_r$  to generate the  $(N+1)$ -electron single determinant  $|^{N+1}\Psi^r\rangle = |\chi_r\chi_1\chi_2 \dots \chi_N\rangle$ , where the remaining spin orbitals are identical to those in  $|^N\Psi_0\rangle$ . In second quantization, this is done by creating an electron in  $\chi_r$ , so that:

$$|^{N+1}\Psi_c\rangle = a_r^\dagger|^N\Psi_0\rangle \quad (2.22)$$

The electron affinity for this process is equal to:

$$EA = {}^N E_0 - {}^{N+1} E^r \quad (2.23)$$

where  ${}^{N+1} E^r$  is the energy of the single determinant, and is expressed as:

$${}^{N+1} E^r = \langle {}^{N+1}\Psi^r | H | {}^{N+1}\Psi^r \rangle \quad (2.24)$$

Therefore, the electron affinity is equal to:

$$EA = -\langle r|h|r\rangle - \frac{1}{2}\sum_a \langle ar||ar\rangle - \frac{1}{2}\sum_b \langle rb||rb\rangle = -\langle r|h|r\rangle - \sum_b \langle rb||rb\rangle = -\varepsilon_r \quad (2.25)$$

If  $\varepsilon_r$  is negative, the electron affinity is positive (*i.e.*, the  $(N+1)$ -electron single determinant  $|^{N+1}\Psi^r\rangle$  is more stable than the  $N$ -electron state  $|^N\Psi_0\rangle$ ).

Koopmans' theorem generally provides ionization energies in qualitative agreement with the experimental ones (within errors of less than 2 eV), but often bad values for electron affinities<sup>62,63</sup>. The neglect of orbital relaxation and of electron correlation represent the two main sources of error; the former is related to the changes in the Fock operator and HF orbitals when changing the number of electrons in the system, while the latter refers to the validity of building the many-body wave function through a single Slater determinant. By going beyond one-determinantal schemes (*i.e.*, by considering electron correlation effects), further corrections to Koopmans' theorem are obtained: since correlation energies are higher for systems with a high number of electrons, correlation effects determine a cancellation of the relaxation error for ionization energies, but add up to it in the case of electron affinities<sup>64</sup>.



## 2.2 Configuration Interaction (CI)

Configuration Interaction (CI, [65]) is a post-HF linear variational method that allows the solution of the nonrelativistic Schrödinger equation for a quantum chemical multi-electron system within the Born–Oppenheimer approximation<sup>66</sup>. However, due to the high computation time and large memory required, it is limited to relatively small systems.

In order to take into account electron correlation, CI employs a variational wave function that is a linear combination of Slater determinants corresponding to electronic configurations in which electrons are excited to virtual orbitals, *i.e.*, orbitals that are not occupied in the HF reference state. Therefore, such determinants represent excited states; their contribution to the total wave function of the ground state in general decreases with the order of excitation. Within this approach, the term *interaction* therefore indicates the mixing (interaction) of different electronic configurations (states). The central idea of the CI approach is the diagonalization of the  $N$ -electron Hamiltonian in a basis of  $N$ -electron Slater determinants. The exact many-electron wave function is therefore expressed as a linear combination of  $N$ -electron trial functions:

$$|\Psi_{CI}\rangle = c_0|\Psi_{HF}\rangle + \sum_S c_S |\Phi_S\rangle + \sum_D c_D |\Phi_D\rangle + \sum_T c_T |\Phi_T\rangle + \dots = \sum_i c_i |\Phi_i\rangle \quad (2.26)$$

where subscripts S, D, T denote determinants that are singly, doubly, triply, excited, respectively, relative to the HF configuration.

In principle, this approach gives an exact solution of the many-electron problem; if the basis set were complete, the exact energy would be obtained not only for the ground state, but also for all excited states. In practice, it is possible to handle only a finite set of  $N$ -electron trial functions; as a consequence, only upper bounds to the exact energies are obtained. Since the number of  $N$ -electron determinants is huge even in the case of small molecules and one-electron basis sets of moderate size, it is necessary to truncate the trial function and employ only a fraction of all possible  $N$ -electron functions; this can be done by considering that there is no mixing of wave functions with different spin. The linear combination in Eq. (2.26) can be reduced to a basis of states with the same spatial and spin symmetry by considering the symmetry properties of the system; the resulting linear combinations of determinants are called *Configuration State Functions* (CSFs). If the expansion in Eq. (2.26) contains all possible CSFs of the appropriate symmetry, it represents a full CI procedure which provides an exact solution of the electronic Schrödinger equation in the given basis. Given the trial function of Eq. (2.26), the corresponding energies can be obtained by means of the linear variational method. This can be done by finding the eigenvalues of the matrix

representation of the Hamiltonian in the basis of the  $N$ -electron functions. In particular, the lowest eigenvalue is an upper bound to the ground state energy of the system, while the higher eigenvalues represent upper bounds to the excited state energies. The difference between the lowest eigenvalue and the HF energy calculated within the same one-electron basis corresponds to the *basis set correlation energy*. The latter approaches the exact correlation energy as the one-electron basis set approaches completeness; however, the basis set correlation energy is exact only within the subspace spanned by the one-electron basis.

## 2.3 Multi-Configuration Self-Consistent Field (MCSCF)

The *Multi-Configuration Self-Consistent Field* (MCSCF, [67]) is a quantum chemistry method used to describe the electronic structure of chemical systems for which a single-electron configuration is not adequate; in particular, the HF energy is corrected by including the static part of the correlation (also known as *near degeneracy effect*). The MCSCF method employs a linear combination of CSFs to approximate the exact electronic wave function of the system. The set of coefficients in front of the determinants in Eq. (2.26), and the MOs employed for the construction of the determinants are both variationally optimized to get the total electronic wave function with the lowest possible energy. This method can be considered as a combination between CI (where only the expansion coefficients are optimized) and HF (where there is only one determinant, but orthonormal MOs are varied). The MCSCF wave function therefore corresponds to a truncated CI expansion, as follows:

$$|\Psi_{MCSCF}\rangle = \sum_I c_I |\Psi_I\rangle \quad (2.27)$$

where a single determinant (or a linear combination of determinants, a CSF) can be used to build  $\Psi_I$ . If only one determinant is included in the expansion, for a closed-shell system, the MCSCF corresponds to the HF method. However, the general equations to solve to get the MCSCF wave function are much more involved than Roothan's restricted HF equations.

In MCSCF, the choice of the most important configurations to describe the property of interest is fundamental, and several implementations of this approach based on different choices of  $\Psi_I$  and the optimization methods are nowadays available. Among them, the CASSCF, which will be outlined in the next paragraph, can be mentioned.

## 2.4 Complete Active Space Self-Consistent Field (CASSCF)

The *Complete Active Space Self-Consistent Field* (CASSCF) is a specific type of MCSCF method which has been put forward by Roos *et al.* .<sup>68</sup> This method is frequently employed in the study of chemical reactions, and is able to provide accurate results for potential energy surfaces; therefore, it is also used as starting point for higher-level multireference methods.

Within the CASSCF method, the number of determinants or CSFs employed in the CI expansion is reduced by dividing the orbitals into two subspaces. In the first subspace (*inactive space*), all orbitals are doubly occupied; these orbitals are not expected to contribute to near-degeneracy correlation effects (*e.g.*, core orbitals). In the second subspace (*active space*), a full-CI expansion is considered by including electrons and orbitals that contribute most to the multireference character of the system. The inactive orbitals have occupation numbers exactly equal to 2, while the occupation numbers of the active orbitals vary between 0 and 2. The total electronic wave function is therefore expressed as a linear combination of all the configurations that satisfy the given spin and space symmetry requirements in the  $N$ -electron space, and is complete in the configurational space spanned by the active orbitals.

The choice of the correct active space is crucial, and an *a priori* knowledge of the investigated system is required. The partitioning between active and inactive orbitals is determined by the target system and the properties to study. Moreover, one has to consider the effective number of electrons and orbitals that can be included in the active space, since the CI expansion increases exponentially with respect to the number of electrons and orbitals. The approach is generally limited to a maximum of 15 active electrons in 15 orbitals; however, this is strongly dependent on the memory and disk storage in the available software packages. The outlined procedure can be indicated as  $[n,m]$ -CASSCF, where  $n$  electrons are distributed in all possible ways in  $m$  orbitals.

The main advantage of this method is that it can be applied to both excited states and the ground state in a single framework. Furthermore, it is well defined on the whole potential energy surface of a chemical reaction, if an appropriate active space is employed. However, it cannot be applied to systems which required a larger active space due to the high computational costs. In particular, there is a factorial dependence on both the number of active electrons and that of active orbitals, which determines the scaling of the method with the size of the system.

The number of Slater determinants in a full CI calculation is equal to:

$$n_{Slater} = \binom{M}{N_\alpha} \binom{M}{N_\beta} \quad (2.28)$$

where  $M$  denotes the number of active orbitals,  $N_\alpha$  and  $N_\beta$  are the numbers of active electrons with spin  $\alpha$  and  $\beta$ , respectively, while the quantities in the parentheses represent the binomial coefficients:

$$\binom{a}{b} = \frac{a!}{b!(a-b)!} \quad (2.29)$$

If spin adapted CSFs are employed, the number of configurations is given by the *Weyl-Robinson formula*<sup>67</sup>, as follows:

$$n_{CSFs} = \frac{(2S+1)}{M+1} \binom{M+1}{M-N/2-S} \binom{M+1}{N/2-S} \quad (2.30)$$

where  $S$  is the total spin and  $N$  the total number of active electrons.

## 2.5 Density Functional Theory (DFT)

Density Functional Theory (DFT) is a well-consolidated quantum-mechanical method, adopted in computational physics and chemistry for the description of the electronic structure of many-body systems (*e.g.*, atoms, molecules and solids).<sup>69</sup> In the last few decades, DFT has greatly expanded the application of quantum-mechanical principles in the investigation of complex systems.

The complete neglect of the electronic correlation in HF methods is overcome in DFT with no increase in computational cost. Indeed, DFT is able to provide a better description of the multielectronic structure at a lower computational cost; in particular, DFT has proven to give very accurate results in the computation of geometries, interatomic distances and vibrational frequencies. In DFT, the electronic structure of atoms, molecules and solids is rigorously described by the electronic density instead of the multielectronic wave function. The energy of the system is a functional of the electronic density. Unlike the wave function, which is a purely mathematical construct, the electronic density is a physical observable. Furthermore, the multielectronic wave function depends on  $4N$  variables (*i.e.*, 3 spatial variables for each electron, and a spin variable), while the electronic density is a function of only three variables (*i.e.*,  $x$ ,  $y$ , and  $z$ ). Considering the computational complexity, another strength of DFT is that it grows much less quickly than *ab initio* methods with the system's dimension, such as HF and post-HF. Indeed, given  $N$  the number of

atoms, DFT scales as  $N^3$ , whereas the simple HF method scales as  $N^4$ ; as a result, DFT represents the method of choice when dealing with very large systems.

### 2.5.1 The electronic Hamiltonian

The electronic structure of a molecular system with  $N$  interacting electrons is determined through the solution of the time-independent electronic Schrödinger equation by assuming the Born-Oppenheimer approximation<sup>66</sup>:

$$H\Psi = E\Psi, \quad (2.31)$$

where  $\Psi$  is the wave function describing the state of the system and depending on electronic coordinates,  $\Psi = \Psi(x_1, x_2, \dots, x_n)$ ,  $E$  is the associated total energy, and  $H$  is the electronic Hamiltonian operator. The latter can be expressed as follows:

$$H = T_e + V_{ee} + V_{ext} = \sum_{i=1}^N \left( -\frac{1}{2} \nabla_i^2 \right) + \sum_{i<j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i=1}^N v_{ext}(\mathbf{r}_i), \quad (2.32)$$

where the first term represents the kinetic energy of the electrons, the second is the electron-electron repulsion energy, and the last term is the external potential (*i.e.*, the attractive potential generated by the nuclei). It is necessary to point out that the exact solution of the Schrödinger equation is hampered by the presence of the two-body electronic potential term.

The eigenvalue equation (Eq. 2.31) can be reformulated by calculating the expectation value of the total energy, as follows:

$$\begin{aligned} E &= \langle \Psi | H | \Psi \rangle = \langle \Psi | T + V_{ee} + V_{ext} | \Psi \rangle = \\ &= -\frac{1}{2} \int \nabla_1^2 \frac{\rho_1(\mathbf{r}'_1, \mathbf{r}_1)}{|\mathbf{r}'_1 - \mathbf{r}_1|} \Big|_{\mathbf{r}'_1 = \mathbf{r}_1} d\mathbf{r}_1 + \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 + \int v_{ext}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \end{aligned} \quad (2.33)$$

where  $\rho(\mathbf{r})$  denotes the electronic density, and represents the probability of finding an electron of arbitrary spin within a volume element  $d\mathbf{r}$ , as follows:

$$\rho(\mathbf{r}) = N \int \dots \int |\Psi(x_1, x_2, \dots, x_N)|^2 ds_1 dx_2 \dots dx_N, \quad (2.34)$$

while  $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$  is the two-particle density function given by:

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{N(N-1)}{2} \int \dots \int |\Psi(x_1, x_2, \dots, x_N)|^2 ds_1 ds_2 dx_3 \dots dx_N \quad (2.35)$$

which represents the probability of finding two electrons simultaneously within the volume elements  $d\mathbf{r}_1$  and  $d\mathbf{r}_2$ . In particular, if  $\Psi$  is normalized,  $\rho(\mathbf{r})$  gives the number of electrons by quadrature, as follows:

$$\int \rho(\mathbf{r}) d\mathbf{r} = N. \quad (2.36)$$

If the wave function is an anti-symmetric product of  $\phi_i(\mathbf{r})$  mono-electronic functions,  $\rho(\mathbf{r})$  is the sum of the probability densities (*i.e.*, probabilities per unit of volume) associated with each  $\phi_i(\mathbf{r})$ :

$$\rho(\mathbf{r}) = \sum_i n_i |\phi_i(\mathbf{r})|^2 \quad (2.37)$$

## 2.5.2 Functionals

The fundamental mathematics concept of DFT is the *functional*. A *functional* is a function whose independent variables are in turn functions; therefore, the domain of a functional is a set of functions, and the functional is an application from the space of functions to the space of real or complex numbers. In mathematical notation:

$$y = F[f(x)]. \quad (2.38)$$

The value of  $y$  is therefore in and of itself dependent on another function. In particular, the total energy of a system,  $E = \langle \Psi | \hat{H} | \Psi \rangle$  is a functional of the electronic density:

$$E = E[\rho]. \quad (2.39)$$

Likewise, the number of electrons  $N = N[\rho]$ :

$$N = \int \rho(\mathbf{r}) d\mathbf{r}, \quad (2.40)$$

where the last equation is a consequence of the normalization of the wave function,  $\Psi$ .

The variation of a functional is given by:

$$F[f + \delta f] - F[f] = \int \frac{\delta F}{\delta f(x)} \delta f(x) dx + O(\delta f)^2 \quad (2.41)$$

where  $\delta f$  represents an infinitesimally small function, and  $O(\delta f)^2$  collects all terms proportional to the higher powers of  $\delta f$ .

## 2.5.3 The Thomas-Fermi model

A variant of DFT was initially conceived by Llewellyn Thomas and Enrico Fermi in 1927,<sup>70,71</sup> immediately following the foundation of quantum mechanics and three years before the devising of the HF theory. Thomas and Fermi were able to use statistical considerations to uniformly distribute electrons in atoms through some approximations. The most important result of this model is that, for the first time, the ground state energy is a functional of the electronic density and not of the wave function.

The electronic density is expressed as function of the Fermi momentum,  $p_f$ . For the highest energy one-electron occupied level,  $p_f$  is defined as follows:

$$p_f = \hbar k_f, \quad (2.42)$$

where  $k_f$  is the Fermi wave vector and is equal to  $\left(\frac{3\pi^2 N}{V_f}\right)^{\frac{1}{3}}$ , in which  $N$  is the number of electrons, and  $V_f$  is the spherical momentum space volume.

The Fermi energy (*i.e.*, the energy of the highest occupied level in the uniform electron gas) is equal to  $\varepsilon_f = \frac{\hbar^2 k_f^2}{2m}$ ; therefore, the electronic density  $\rho_0$  for a uniform electron gas is given by:

$$\rho_0 = \frac{p_f^3}{3\pi^2 \hbar^3} \quad (2.43)$$

In the Thomas-Fermi model of the atom, the functional of the total energy is given by:

$$E_{TF}[\rho(\mathbf{r})] = C_{TF} \int \rho^{\frac{5}{3}}(\mathbf{r}) d\mathbf{r} - Z \int \frac{\rho(\mathbf{r})}{r} d\mathbf{r} + \frac{1}{2} \iint \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \quad (2.44)$$

with  $C_{TF} = \frac{3}{10} (3\pi^2)^{\frac{2}{3}} = 2.871$  (in atomic units).

The first term of Eq. (2.44) corresponds to the kinetic energy of electrons, the second one to the electrons-nucleus attraction, and the last one to the repulsion among electron pairs. It is necessary to point out that Eq. (2.44) represents an approximate form of the rigorous expression for the electronic energy, in which only the stabilizing (nucleus-electron) and destabilizing inter-electronic classical interactions are considered.

Finally, some important considerations on the validity of this model are in order. First, it is exact only for the uniform electron gas in the limit of an atomic number  $Z \rightarrow \infty$ , as it gives a rational description only of the charge density for heavy atoms.<sup>72</sup> Furthermore, it cannot provide accurate results in atomic and molecular systems, since it is not able to self-consistently represent the atomic shell structure, and gives errors of about 10% in the resulting energies. Finally, the existence of any chemical bonds cannot be shown, since, for an aggregate of atoms, it states that the minimum energy is given by nuclei at infinite distance.<sup>73</sup>

## 2.5.4 The birth of modern DFT: Hohenberg and Kohn theorems

In the 1960s, Hohenberg and Kohn (HK) developed the modern version of DFT by using the concepts of the Thomas-Fermi model, and were able to demonstrate that the latter is an approximate form of an exact theory.<sup>74</sup> Originally, the HK theorems held only for non-degenerate ground states in the absence of a magnetic field, but then have been generalized by M. Levi and J.P. Perdew (*Constrained Search Formalism*, [75]).

To outline HK theorems, one can consider a system of  $N$  electrons and a fixed molecular geometry. In the non-relativistic case and adopting the Born-Oppenheimer approximation, the electronic Hamiltonian is given by:

$$H = \sum_{i=1}^N \left( -\frac{1}{2} \nabla_i^2 \right) + \sum_{i < j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i=1}^N v_{ext}(\mathbf{r}_i) \quad (2.45)$$

where the first term represents the electronic kinetic energy operator ( $T_e$ ), the second one is the repulsive potential energy between electrons ( $V_{ee}$ ), and  $v_{ext}$  is the external potential.

**First HK theorem:** In a finite interacting  $N$ -electron system, the external potential, and therefore the total energy, is a unique functional of the ground state density, up to an arbitrary additive constant.

It follows that there is a *one-to-one* correspondence between the external potential and the ground state density.<sup>76</sup> It is worth noting that this theorem does not provide any information neither on the analytic form of the energy functional nor on practical ways to get the ground state electronic density.

**Proof:** One considers two potentials, namely,  $v(\mathbf{r})$  and  $v'(\mathbf{r})$ , which give the same ground state density  $\rho_{GS}$ , and which differ by more than a constant (*i.e.*,  $v(\mathbf{r}) - v'(\mathbf{r}) \neq \text{constant}$ ). For both potentials, the following equality holds:

$$\int \rho_{GS}(\mathbf{r}) d\mathbf{r} = N \quad (2.46)$$

Hence, one obtains:

$$H = T_e + V_{ee} + \sum_i v(\mathbf{r}_i),$$

$$H' = T_e + V_{ee} + \sum_i v'(\mathbf{r}_i), \quad (2.47)$$

$$H - H' = \sum_i v(\mathbf{r}_i) - \sum_i v'(\mathbf{r}_i) \neq \text{constant}.$$

If one assumes that  $\Psi$  and  $\Psi'$  are the ground state wave functions of  $H$  and  $H'$ , respectively, and are both normalized:

$$\begin{aligned} E_0 &= \langle \Psi | H | \Psi \rangle < \langle \Psi' | H' | \Psi' \rangle = \langle \Psi' | H' - \sum_i (v(\mathbf{r}_i) - v'(\mathbf{r}_i)) | \Psi' \rangle = \\ &= E_0' + \int \rho(\mathbf{r}) [v(\mathbf{r}) - v'(\mathbf{r})] d\mathbf{r} \end{aligned} \quad (2.48)$$

By analogy:

$$\begin{aligned} E_0' &= \langle \Psi' | H' | \Psi' \rangle < \langle \Psi | H' | \Psi \rangle = \langle \Psi | H - \sum_i (v(\mathbf{r}_i) - v'(\mathbf{r}_i)) | \Psi \rangle = \\ &= E_0 + \int \rho(\mathbf{r}) [v'(\mathbf{r}) - v(\mathbf{r})] d\mathbf{r} \end{aligned} \quad (2.49)$$



By summing the last two equations, one obtains the following contradiction:

$$E_0 + E_0' < E_0 + E_0' \quad (2.50)$$

As both  $T_e$  and  $V_{ee}$  are fixed by the ground state electronic density  $\rho_{GS}$ , it follows that the Hamiltonian is a functional of the ground state density  $\rho_{GS}(\mathbf{r})$ , and, therefore, all eigenstates of the Schrödinger equation are functionals of the density (*i.e.*,  $\Psi[\rho(\mathbf{r})]$ ). Hence, the energy can be expressed as a functional of the density:

$$E[\rho] = \langle \Psi[\rho] | T_e + V_{ee} + V_{ext} | \Psi[\rho] \rangle \quad (2.51)$$

**Second HK theorem** (variational principle of the energy): the ground state energy  $E_{GS}$  is always smaller or equal to the energy functional evaluated with a trial electronic density  $\rho(\mathbf{r})$ :

$$E[\rho(\mathbf{r})] \geq E_{GS} \quad (2.52)$$

with  $[\rho(\mathbf{r})] \geq 0$  and  $\int \rho(\mathbf{r}) d(\mathbf{r}) = N$

In particular, the equality is verified when  $\rho(\mathbf{r})$  is the exact electronic density of the ground state.

**Proof.** One considers that the energy of the ground state is uniquely defined by the ground state density  $\rho_{GS}(\mathbf{r})$ :

$$E_{GS} = E[\rho_{GS}] = \langle \Psi_{GS} | H_{GS} | \Psi_{GS} \rangle \quad (2.53)$$

Furthermore, according to the variational principle, a different density  $\rho(\mathbf{r})$  necessary gives a higher energy, as follows:

$$E_{GS} = E[\rho_{GS}] = \langle \Psi_{GS} | H_{GS} | \Psi_{GS} \rangle < \langle \Psi | H_{GS} | \Psi \rangle = E[\rho(\mathbf{r})] \quad (2.54)$$

The exact density  $\rho_{GS}(\mathbf{r})$  of the interacting  $N$ -electron system can be then calculated by minimizing the energy functional  $E[\rho]$ . The minimization problem can be solved by introducing a constraint, since the system has a constant number of electrons ( $N$ ). A *Lagrangian function* can be hence defined:

$$L = E[\rho] - \mu(\int \rho(\mathbf{r}) d\mathbf{r} - N), \quad (2.55)$$

where  $\mu$  denotes the Lagrangian multiplier.

The minimum condition of the energy functional corresponds to zero functional derivative (*Euler equation*):

$$\frac{\delta}{\delta \rho} [E[\rho] - \mu(\int \rho(\mathbf{r}) d\mathbf{r} - N)] = 0. \quad (2.56)$$

Eq. (2.51) can be rewritten by defining a functional for each term of the Hamiltonian:

$$E[\rho] = F[\rho] + \int v_{ext}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}, \quad (2.57)$$

where  $v_{ext}(\mathbf{r})$  is the external potential, while  $F[\rho]$  represents the *Hohenberg and Kohn functional*, that is a universal functional that does not depend on the external potential, but only on the number of electrons  $N$ .  $F[\rho]$  is given by the sum of two different contributions:

$$F[\rho] = \langle \psi | T_e + V_{ee} | \psi \rangle = T[\rho] + V_{ee}[\rho], \quad (2.58)$$

where  $V_{ee}[\rho] = V_{XC}[\rho] + J[\rho]$ .  $V_{XC}[\rho]$  is the exchange-correlation potential, while  $J[\rho]$  represents the classic Coulomb repulsion energy between two electronic densities:

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2. \quad (2.59)$$

Eq. (2.56) can be therefore rewritten in the following way:

$$\frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} + V_{ext}(\mathbf{r}) = \mu. \quad (2.60)$$

It is necessary to point out that the HK functional is defined only for the *v-representable* ground state densities  $\rho(\mathbf{r})$ . In particular, a *v-representable* density is an electronic density obtained from a antisymmetric ground state wave function of an Hamiltonian of the form of Eq. (2.45). The *N-representability* of an electronic density  $\rho$ , which is a weaker condition compared to the *v-representability*, is satisfied by the following conditions:

$$\rho(\mathbf{r}) \geq 0, \int \rho(\mathbf{r}) d\mathbf{r} = N \text{ and } \int \left| \nabla \rho(\mathbf{r})^{\frac{1}{2}} \right|^2 < \infty. \quad (2.61)$$

The limitation to *v-representable* description is removed by the constraint search formalism<sup>75</sup>, which adopts the *Rayleigh-Ritz* variational principle to determine the ground state energy associated with the electronic Hamiltonian of the *N-electron* system (Eq. 2.45). Despite the fundamental formal and conceptual role of the HK functional with the constraint search formalism in DFT, the HK functional is unknown. Furthermore, this approach does not give any operative strategy to be employed in practice, although it is essential for a formal justification of the theory. In particular, the second HK theorem is only an existence theorem from which one cannot extract any information on the construction of the ground state energy functional. These limitations can be overcome by the Kohn-Sham approach<sup>77</sup> that will be treated in the next paragraph.

## 2.5.5 The Kohn-Sham approach

The computational scheme of DFT is determined by the Kohn-Sham (KS) equations [Kohn, L.J. Sham, *Phys. Rev.* **1965**, *140*, A1133-A1138]. In particular, in 1965 Kohn and Sham proposed a powerful approach to indirectly treat the kinetic energy functional, introducing orbitals into the problem to determine the kinetic energy at a good approximation.

Within this approach, along with the “real” system of interacting electrons, one considers a “fictitious” system of  $N$  *non-interacting* electrons, but with the same electronic density of that of the

“real” system. Therefore, the intractable many-body problem of interacting electrons in a static external potential is reduced to a tractable problem of non-interacting electrons moving in an effective potential.

In the “fictitious” system, the KS Hamiltonian (*Kohn-Sham Hamiltonian*) is given by:

$$H_{KS} = \sum_{i=1}^N \left[ -\frac{1}{2} \nabla_i^2 + v_{KS}(\mathbf{r}_i) \right] = \sum_{i=1}^N h_{KS}(i) \quad (2.62)$$

where  $v_{KS}$  is the KS potential and must be determined. It is a local mono-electronic potential that has to be chosen so that the density of the KS system is equal to the electronic density of the interacting one. The energy functional is expressed as:

$$E_{KS}[\rho] = T_0[\rho] + \int \rho(\mathbf{r}) v_{KS}(\mathbf{r}), \quad (2.63)$$

where  $T_0[\rho]$  is the non-interacting kinetic energy functional, and is defined as:

$$T_0[\rho] = -\frac{1}{2} \sum_{i=1}^N \langle \phi_i | \nabla_i^2 | \phi_i \rangle. \quad (2.64)$$

Hence, Eq. (2.56) becomes:

$$\frac{\delta E_{KS}[\rho]}{\delta \rho(\mathbf{r})} = \frac{\delta T_0[\rho]}{\delta \rho(\mathbf{r})} + v_{KS}(\mathbf{r}) = \mu \quad (2.65)$$

providing the exact density of the ground state in the non-interacting system.

The  $N$ -electron ground state wave function can be constructed through a single Slater determinant, with  $\phi_i(\mathbf{r})$  orbitals satisfying the time-independent Schrödinger equation:

$$h_{KS} \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}) \quad (2.66)$$

$$\text{where } h_{KS} = -\frac{1}{2} \nabla^2 + v_{KS}. \quad (2.67)$$

Eq. (2.66) represents the KS mono-electronic equations, whose solutions are the  $\phi_i(\mathbf{r})$  orbitals (*Kohn-Sham orbitals*) and the  $\varepsilon_i$  orbital energies. According to the KS equations, electrons in atoms, molecules and solids are independent particles that move in the effective potential  $v_{KS}$ .

The KS wave functions are functionals of the density  $\phi_i[\rho(\mathbf{r})]$ :

$$\rho_{GS}^{KS}(\mathbf{r}) = \sum_{i=1}^N |\phi_i(\mathbf{r})|^2 \quad (2.68)$$

Let's consider now the case of the real (interacting) system; the energy functional can be written as follows:

$$E[\rho] = T_0[\rho] + J[\rho] + E_{xc}[\rho] + \int v_{ext}(\mathbf{r}) \rho(\mathbf{r}) \quad (2.69)$$

where  $E_{xc}[\rho]$  is the *exchange and correlation functional* which is defined as:

$$E_{xc}[\rho] = T[\rho] - T_0[\rho] + V_{ee}[\rho] - J[\rho] \quad (2.70)$$

Performing the functional derivatives in Eq. (2.56), one obtains the minimum condition for the real system:

$$\frac{\delta T_0}{\delta \rho} + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{XC}}{\delta \rho(\mathbf{r})} + v_{ext}(\mathbf{r}) = \mu \quad (2.71)$$

Requiring that the two systems have the same electronic density  $\rho(\mathbf{r})$ , the following expression for  $v_{KS}(\mathbf{r})$  is found:

$$v_{KS}(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{XC}}{\delta \rho(\mathbf{r})} + v_{ext}(\mathbf{r}) \quad (2.72)$$

where the first term is the electrostatic potential generated by the electronic density of the system, the second one is the *exchange-correlation potential*, which represents the difference between the exact kinetic energy functional and that of the non-interacting kinetic energy functional together with all non classical terms of the electron-electron interaction.

The ground state density of the interacting system can be hence determined through the solution of the time-independent Schrödinger equation by employing Eq. (2.66):

$$\rho_{GS}(\mathbf{r}) = \sum_i n_i |\phi_i(\mathbf{r})|^2. \quad (2.73)$$

As  $v_{KS}$  depends on the ground state density  $\rho$ , which in turns depends on the mono-electronic wave functions  $\phi_i$ , according to Eq. (2.68), an iterative process is necessary to solve the KS equations; this process is denoted as the *Self Consistent Field* (SCF). Within the latter, a trial density  $\rho$  is used to determine the trial solutions of the KS equations, which are employed to generate a new electronic probability density that triggers the iterative procedure until convergence is reached, *i.e.*, until the difference between the energy calculated at the current cycle and the energy calculated at the previous cycle is lower than a predefined threshold. In particular, the commutator of the KS matrix and the *P-matrix* from which the KS matrix is extracted must be equal to zero for an exact solution.

## 2.5.6 Exchange and correlation potentials

In order to determine uniquely the KS Hamiltonian operator, an explicit form of the exchange-correlation (xc) potential has to be defined explicitly; as observed before, the latter is a term of the local KS effective potential (Eq. 2.72), and include all non-classical contributions to the Coulombic interaction energy. However, the exact form of the xc energy functional is unknown; therefore, the xc potential must be approximated. In literature, several approximations are available; the choice of the approximation is based on the system and properties to investigate, and strongly affects the accuracy of the results. The accuracy achievable by the chosen xc potential is directly related to the

quality of the approximation adopted for the so-called *exchange-correlation* (xc) *hole*, which is a pair correlation function that includes non-classical effects, and is made of two contributions:

$$h_{xc}(\mathbf{r}_1, \mathbf{r}_2) = h_x^{s_1=s_2}(\mathbf{r}_1, \mathbf{r}_2) + h_c^{s_1, s_2}(\mathbf{r}_1, \mathbf{r}_2) \quad (2.74)$$

where the exchange part is the *Fermi hole*, which states the Pauli's exclusion principle, and affects electrons with the same spin value (s), while the dynamical correlation gives rise to the *Coulomb hole*, which is due to electrostatic effects, that affect all electrons independently from their spin value. However, only the overall expression of the hole has a physical meaning. Since the exchange and the correlation keep the electrons apart, the exchange and the correlation contributions can be interpreted in terms of a hole surrounding each electron, and preventing other electrons to approach it. The two-particle density function  $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$  in Eq. (2.35) can be written as:

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \rho(\mathbf{r}_1)\rho(\mathbf{r}_2)[1 + h_{xc}(\mathbf{r}_1, \mathbf{r}_2)] \quad (2.75)$$

This represents a conditional probability, and is known as *exchange-correlation hole*.

The expectation value of the inter-electronic repulsive potential  $V_{ee}$  hence becomes:

$$\langle V_{ee} \rangle = \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 + \frac{1}{2} \iint \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)h_{xc}(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 = J[\rho] + v_{xc}(\mathbf{r}_1) \quad (2.76)$$

where the first term represents the classic interaction between two electronic densities, and also contains a self-interacting term, while the second one counts for the interaction between the electronic density and the xc hole, introducing the correlation effects and the correction to the self-interaction.

### 2.5.6.1 Local Density Approximation (LDA)

The *Local Density Approximation* (LDA, [78]) is the oldest approximation to the xc potential, and is based upon the exact exchange energy for a uniform electron gas. It consists in partitioning the whole 3D space into small volume elements in which the electronic density can be assumed to be constant; the xc energy functional is assumed to be only dependent on the value of the electronic density at each point of the space.

The LDA xc energy functional of an inhomogeneous system is defined as the integral over the xc energy per particle in the homogeneous electron gas at a given density  $\rho(\mathbf{r})$ :

$$E_{XC}^{LDA}[\rho(\mathbf{r})] = \int \varepsilon_{XC}^{LDA}[\rho(\mathbf{r})]\rho(\mathbf{r})d\mathbf{r}. \quad (2.77)$$

The energy per particle of an homogeneous gas can be divided into two terms:

$$\varepsilon_{xc} = \varepsilon_x + \varepsilon_c, \quad (2.78)$$

where the exchange energy contribution,  $\varepsilon_x$ , is defined as:

$$E_X^{\text{LDA}}[\rho(\mathbf{r})] = -\frac{3}{4} \left(\frac{3}{\pi}\right)^{\frac{1}{3}} \int \rho(\mathbf{r})^{\frac{4}{3}} d\mathbf{r}, \quad (2.79)$$

while the correlation term  $\varepsilon_c$  can be approximated by means of accurate parametrizations, *e.g.*, analytically by interpolating quantum Monte Carlo results on an homogeneous electron gas<sup>79-81</sup>.

LDA represents the most accurate choice for systems whose electronic density does not undergo significant variations in space. In order to evaluate if LDA is a good approximation for a system, one has to consider the local Fermi wave vector  $K_f(\mathbf{r})$ . In particular, LDA represents a good choice if the following condition is verified<sup>78</sup>:

$$\frac{|\nabla\rho(\mathbf{r})|}{\rho(\mathbf{r})} \geq K_f(\mathbf{r}). \quad (2.80)$$

Being based on the homogeneous electron liquid, which is a real physical system (and therefore satisfies a lot of exact conditions), LDA works well in the description of several material properties, such as lattice constants of solids, dipole moments, vibrational frequencies, Fermi surfaces of bulk materials. In particular, it gives atomic and molecular total ground state energies values within about 1-5% of the exact ones, and molecular equilibrium distances and geometries within about 3%.<sup>76</sup> However, this approximation tends to overestimate polarizabilities, and to underestimate electron affinities and IPs. Moreover, KS eigenvalues result to be too low in magnitude, due to the fact that the xc potential does not go to zero as  $-1/r$  at large distances, but exponentially fast. In the last decades, great efforts have been made to develop suitable corrections to the LDA approximation.

### 2.5.6.2 General Gradient Approximation (GGA)

*General Gradient Approximations* (GGAs) are used for systems characterized by fast spatial variations of the electronic density. The GGA takes into account the inhomogeneous character of the electronic density by considering its gradient; the xc energy functional can be calculated as follows:

$$E_{XC}^{\text{GGA}}[\rho(\mathbf{r})] = \int \rho(\mathbf{r}) \varepsilon_{XC}^{\text{GGA}}(\rho, \nabla\rho) d\mathbf{r}. \quad (2.81)$$

GGA functionals are *non-local*, since the xc energy at a given point ( $\varepsilon_{xc}(\mathbf{r})$ ) is not uniquely defined by the electronic density in that point ( $\rho(\mathbf{r})$ ), but also by the electronic density in the

neighboring points ( $\rho(\mathbf{r} + \Delta\mathbf{r})$ ). As the gradient-dependent terms are usually small for both atoms and molecules, the GGA functional is also known as *gradient corrected functional*.

GGA potentials are classified according to the degree through which they follow the exact asymptotic behavior at large distances (*i.e.*, as  $-1/r$ ). The LB94 potential by van Leeuwen and Baerends<sup>82</sup> is one of the most popular GGA functionals, and is designed to simulate the Coulombic decay of the exact xc potential. Other famous GGA functionals include the B88 exchange functional<sup>83</sup>, the LYP correlation functional<sup>84</sup>, and the PBE functional<sup>85</sup>. The PW86x potential by Perdew and Wang<sup>86</sup> is one of the first GGA functionals that gives more accurate results than the LDA parents.

The most promising routes that allow to systematically improve the GGA functional are identified in *empirical* and *non-empirical* functionals. The former are obtained by fitting free parameters to reproduce exact properties of several molecules or atoms; the M06-L functional by Truhlar and collaborators<sup>87</sup> is included in this category. The latter are instead derived through approximate xc holes, and free parameters are chosen to satisfy mathematical conditions, *e.g.*, asymptotic behavior, behavior of energy at fractional occupancies, scaling, bounds, etc.

To improve GGA, a reduced density gradient  $s(\mathbf{r})$  can be used to build expansions of density functionals in power of gradients (based on Eq. 2.80). The reduced density gradient is expressed as follows:

$$s(\mathbf{r}) = \frac{|\nabla\rho(\mathbf{r})|}{2\rho(\mathbf{r})K_f(\mathbf{r})}. \quad (2.82)$$

By using a *Gradient Expansion Approximation* (GEA), one can obtain the following expression for  $E_{xc}$ :

$$E_{xc}^{GEA}[\rho] = \int \left( \varepsilon_{xc}(\rho) + C_{xc}^{(2)}(\rho)s^2 + \dots \right) d\mathbf{r}. \quad (2.83)$$

Another family of xc functionals is formed by the so-called *meta-GGA* functionals, which use higher derivatives of the electronic density to approximate the density of the xc energy;<sup>88</sup> their formula also includes a contribution related to the positive kinetic energy density of the occupied KS orbitals.<sup>89</sup> Two more classes of xc functionals involving more complicated and accurate approximations can be mentioned, namely, the *hyper-GGA* and the *generalized random phase approximation* functionals. In the former, the exact exchange energy density is included,<sup>90</sup> while the latter take into account the unoccupied orbitals.<sup>91</sup> It has been found out that more accurate results

can be generally obtained by increasing the approximations' complexity at computational costs that modestly increase from the simple LDA to the more complex meta-GGA.<sup>92</sup>

### 2.5.6.3 Hybrid exchange and correlation potentials

Hybrid xc potentials were introduced by A. Becke in 1993<sup>93</sup>; these potentials incorporate a part of exact exchange from the HF theory, and employ KS orbitals, rather than density, to compute the exact exchange energy functional. This hybrid formulation is able to provide a good description of short range electron-electron interactions, and, as a consequence, very accurate results for short-range properties; reliable results can be also obtained for vibrational frequencies, atomization energies, and bond lengths. There are two kinds of hybrid functionals, which can be classified as purely theoretical or semi-empirical, depending on whether their parameters are fitted with experimental data or not. The hybrid functional known as B3 (Becke's three parameter exchange functional) is nowadays one of the most commonly used functionals, usually in combination with either LYP or PW91 correlation functionals. The B3LYP xc potential<sup>60</sup>, which is an example of semi-empirical functional, possesses a 15% of HF exchange, and is expressed as follows:

$$E_{XC}^{B3LYP} = E_x^{LDA}[\rho] + \alpha_0(E_x^{HF} - E_x^{LDA}) + \alpha_x \Delta E_x^{B88} + \alpha_c E_c^{LYP} + (1 - \alpha_c) E_c^{VWN} \quad (2.84)$$

where  $\alpha_0$ ,  $\alpha_x$ ,  $\alpha_c$  are the three parameters which define the potentials, and are approximately equal to 0.2, 0.7, 0.8, respectively.  $E_x^{B88}$  is the exchange functional *Becke 88* [83], while  $E_c^{LYP}$  is the *Lee, Yang, Parr* correlation functional<sup>94</sup>; these functional use GGA, while  $E_x^{LDA}$  employs LDA.

The xc potentials used in NEXAFS spectra calculations of each system investigated in this thesis work will be reported in the Computational Details section of the attached published papers.

## 2.5.7 The calculation of NEXAFS spectra through DFT: the Transition Potential

### (TP) method

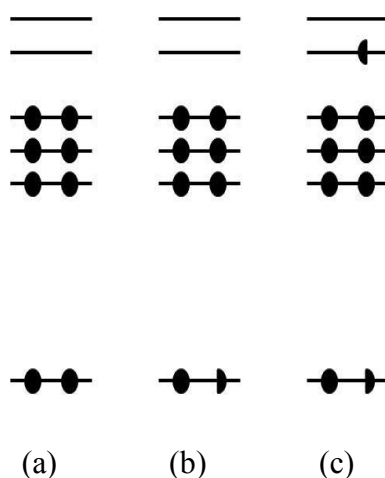
In this thesis work, DFT NEXAFS spectra calculations have been performed through the *Transition Potential* (TP) method developed by J.C. Slater [25], in which relaxation effects are described through second order perturbation theory. Slater analyzed the eigenvalues of excited states, and found out that SCF calculations performed on non-physical states in between the ground state and the excited state of the *N-electron* system provided particularly accurate values for the excitation energies. The main strength of this approach is the inclusion of the relaxation effects (*i.e.*, the modification of the MOs upon the core-hole formation) at an affordable computational cost. Indeed, due to the significant reduction in the screening of the nucleus following the core excitation, the inclusion of relaxation effects is fundamental for a proper description of the phenomenon. This



method has been later generalized by Goscinski *et al.*<sup>95,96</sup> and by Williams *et al.*<sup>97</sup> to take into account higher-order perturbation terms.

In the X-ray absorption process, the excited electron may remain bound either in a Rydberg or in an empty MO; the main problem resides in the way to calculate transition energies for such "two level" cases, involving a core hole and an electron in an "outer" orbital, respectively. In one variant of the TP method, half an electron is removed from the core orbital and added to the upper orbital; transition energies are then calculated as the differences between the orbital energies of the core and virtual orbital involved in the excitation. However, this approach requires an additional calculation for each virtual orbital; to avoid it and perform both bound and continuum state computations with the same potential, one can neglect the contribution of the excited outer shell electron to the molecular potential, and setting the occupation of the core orbital by one half, leaving all virtual MOs unoccupied.

Figure 2.1 schematically shows the MO occupations of the ground state (a) and for two different TP schemes (b, c) that can be employed to compute core-excitation energies and oscillator strengths.



**Figure 2.1** – A schematic representation of the ground state (a) of the TP model used in this thesis work (b) and of another variant of the TP model (c).

Scheme (b) depicts an  $(i)^{-1/2} (a)^0$  electronic configuration, *i.e.*, obtained by removing half an electron from the core orbital  $i$ , leaving all virtual orbitals unoccupied, while in scheme (c) an  $(i)^{-1/2} (a)^{1/2}$  electronic configuration is represented, where half an electron is promoted to a virtual orbital. Scheme (b) has been employed in this thesis work, while scheme (c) is used when calculations of a particular transition energy are required<sup>98,99</sup>.

Computationally, the mono-electronic TP-KS equations are solved:

$$h_{KS}^{TP} \varphi_i^{TP} = \varepsilon_i^{TP} \varphi_i^{TP} \quad \text{with } i = 1, \dots, n, \quad (2.85)$$

in which the  $h_{KS}^{TP}$  Hamiltonian is built through an  $(i)^{-1/2} (a)^0$  or  $(i)^{-1/2} (a)^{1/2}$  electronic configuration.

Excitation energies are obtained as the difference between the energy eigenvalues relative to the final virtual orbital and the core orbital involved in the transition:

$$\Delta E_{i \rightarrow f} = \varepsilon_f^{TP} - \varepsilon_i^{TP}. \quad (2.86)$$

The ionization potentials (IPs) are directly calculated following Koopmans' theorem<sup>57</sup> as the opposite of the TP eigenvalue of the core orbital:

$$IP^{TP} = -\varepsilon_i^{TP} \quad (2.87)$$

Since the TP-DFT approach generally leads to a less attractive potential, the absolute transition energies result to be higher than expected. To account for this error, it is convenient to first compute the IPs within the  $\Delta$ KS ( $\Delta$ SCF Kohn–Sham) scheme, hence allowing a full relaxation of the core hole, and then scaling the TP excitation energies (Eq. 2.86) by the  $\varepsilon_i^{TP} - IP^{\Delta KS}$  energy difference. In particular, two distinct SCF calculations are performed, namely, one for the  $N$ -electron ground state, and one for the  $N-1$ -electron configuration (*i.e.*, the ion). The  $IP^{\Delta KS}$  is given by the difference between the energy of the ion and that relative to the ground state.

Transition intensities are expressed through the dimensionless quantity *oscillator strength*,  $f_{i \rightarrow f}$ . In the case of randomly-oriented gas phase molecules,  $f_{i \rightarrow f}$  is defined as:

$$f_{i \rightarrow f} = \frac{2}{3} n_i \Delta E_{i \rightarrow f} |\langle \varphi_f^{TP} | \boldsymbol{\mu} | \varphi_i^{TP} \rangle|^2 \quad (2.88)$$

involving dipole integrals between the eigenfunctions corresponding to the initial and final states.  $n_i$  is the occupation number of the core orbital in the ground state (normally equal to 2), while  $\Delta E_{i \rightarrow f}$  is the transition energy calculated through the TP method.

When dealing with molecules adsorbed on surfaces (oriented molecules), Eq. (2.88) has to be rewritten as follows:

$$f_{i \rightarrow f} = 2 n_i \Delta E_{i \rightarrow f} |\langle \varphi_f^{TP} | \hat{\mathbf{e}} \cdot \boldsymbol{\mu} | \varphi_i^{TP} \rangle|^2 \quad (2.89)$$

where  $\hat{\mathbf{e}}$  is the polarization vector of the incident radiation.

In the case of XP spectra, transition intensities cannot be calculated by means of the herein reported oscillator strength expressions, since the implicated excited states lay in the electronic continuum. Therefore, the computation of the electronic continuum wave function at a given photoelectron kinetic energy would in principle be required for an accurate calculation of the XPS intensity.

However, since the photon energy used in XP measurements is usually well above the ionization threshold, at good approximation, the XPS intensity for each nonequivalent site can be assumed to be proportional to the number of equivalent centers in the investigated molecule.

## 2.6 Density of states (DOS) calculations

Density of states (DOS) calculations are commonly performed in solid-state and condensed matter physics and chemistry. In particular, partial (projected) DOS (PDOS) calculations are very useful in the case of cluster model used to mimic solid state systems (as those investigated in this thesis work), since they allow a pictorial representation of Mulliken populations relative to the cluster virtual orbitals. Moreover, PDOS calculations are able to provide an accurate characterization of the electronic structure in terms of atomic-like contributions to the virtual orbitals, and determine the character of these contributions (*i.e.*, bonding, nonbonding or antibonding) with respect to a particular bond.

The DOS of a system describes the number of states per energy interval; therefore, it gives an idea about how energy levels are closely packed in a certain system. It is hence represented by a density distribution: a high DOS at a specific energy indicates that there are many states available for occupation, while a DOS of zero means no states at that energy level. At 0K the states are occupied up to the Fermi energy, while, above it, the states are empty.

The number of one-electron levels (*i.e.*, orbitals) in the energy range  $dE$  between  $E$  and  $E+dE$  is expressed by the quantity  $N(E)dE$ . The total density of states (TDOS) at energy  $E$  is given by:

$$N(E) = \sum_i \delta(E - \varepsilon_i) \quad (2.90)$$

where  $\varepsilon_i$  are the one-electron energies.

The number of one-electron states in the energy range  $[E_1, E_2]$  can be written as:

$$N_{E_1, E_2} = \int_{E_1}^{E_2} N(E) dE \quad (2.91)$$

The  $\delta$ -functions in Eq. (2.90) are usually broadened with Lorentzian functions to obtain a graphical representation of Mulliken populations, and, when weighted by proper coefficients describing some properties of the electronic level  $\phi_i$  at the energy  $\varepsilon_i$ , different typologies of DOS are built. When one has to find out if a given atomic function  $\chi_\mu$  (relative to the absorbing atom) contributes strongly to MOs at certain energies, the different one-electron levels may be weighted by using the

percentage  $\chi_\mu$  character; if the latter is determined by the gross populations, the Gross Populations DOS (GPDOS) is obtained:

$$N_\mu(E) = \sum_i GP_{i,\mu} L(E-\varepsilon_i) \quad (2.92)$$

where the index  $i$  runs over the one-electron energy levels,  $\varepsilon_i$  is the energy of the  $i$ -th level,  $L(E-\varepsilon_i)$  is a proper Lorentzian function (of a given FWHM), and  $GP_{i,\mu}$  is the gross population of the  $\chi_\mu$  function in a specific orbital density expressed in terms of the  $P_i$  density matrices:

$$GP_{i,\mu} = \sum_\nu P_{i,\mu\nu} S_{\mu\nu} = \sum_\nu C_{\mu i} C_{\nu i} S_{\mu\nu} \quad (2.93)$$

where  $S_{\mu\nu}$  is the overlap matrix element between the  $\chi_\mu$  and the  $\chi_\nu$  basis functions, and  $C_{\mu i}$  are the coefficients of the  $\chi_\mu$  function in the  $i$ -th  $\phi_i$  MO:

$$\phi_i(\mathbf{r}) = \sum_\mu \chi_\mu(\mathbf{r}) C_{\mu i} \quad (2.94)$$

The gross population  $GP_{i,\mu}$  of a function in a specific orbital density  $|\phi_i(\mathbf{r})|^2$  is therefore associated to the fraction of the orbital density belonging to that function (*i.e.*, percentage  $\chi_\mu$  character of the  $\phi_i$  MO); in particular, the GPDOS assumes large values at energies where  $\chi_\mu$  occurs in the MOs.

If the weight factor is determined by projection of  $\chi_\mu$  onto the orbital  $\phi_i$ , one obtains the PDOS:

$$N_\mu(E) = \sum_i |\langle \chi_\mu | \phi_i \rangle|^2 L(E-\varepsilon_i) \quad (2.95)$$

Finally, if the  $\delta$ -function representing orbital  $\phi_i$  is weighed with the overlap population between  $\chi_\mu$  and  $\chi_\nu$  in  $\phi_i$ , the Overlap Population Density Of States (OPDOS) is obtained:

$$N_{\mu\nu}(E) = \sum_i OP_{i,\mu\nu} L(E-\varepsilon_i) \quad (2.96)$$

In particular, the overlap population is strongly positive, and OPDOS(E) is large and positive around  $E = \varepsilon_i$  when an orbital  $\phi_i$  at energy  $\varepsilon_i$  is strongly bonding between  $\chi_\mu$  and  $\chi_\nu$ , whereas, OPDOS(E) is negative around  $E = \varepsilon_i$  when the orbital  $\phi_i$  is antibonding between  $\chi_\mu$  and  $\chi_\nu$ . The OPDOS(E) has been referred as *coop* (*i.e.*, crystal orbital overlap population) in extended-Hückel solid state calculations performed by R. Hoffmann and coworkers.<sup>100</sup>

## 2.7 Time-Dependent Density Functional Theory (TDDFT)

Although DFT has been extended in different manners for the treatment of the excited states, such as in the case of the TP method previously described, DFT itself provides a precise description only of the electronic structure of the ground state. A rigorous theoretical framework to study excited states is provided by Time-Dependent Density Functional Theory (TDDFT), whose theoretical bases have been provided in 1984 by Runge and Gross by demonstrating the analogous of the HK theorems extended to time-dependent systems. This method has been widely employed in theoretical calculations of excitation energy for both closed<sup>101</sup> and open shell<sup>102</sup> systems. TDDFT currently represents a good compromise between accuracy and computational cost, and has also been implemented in the ADF code<sup>103</sup>. Moreover, TDDFT is considered the state-of-the-art theoretical approach for the computation of core excitation energies in the DFT framework; the configuration mixing introduced beyond the simple KS scheme is particularly relevant in the presence of degenerate core holes, and gives a large intensity redistribution with consequent improvements in the matching with the experimental<sup>31,32</sup>.

When dealing with systems containing heavy atoms, relativistic effects must be also taken into account in order to obtain reliable results. In particular, excitation energies and transition intensities are essentially affected by two kinds of relativistic effects, namely, scalar-relativistic corrections and spin-orbit (SO) coupling. The first ones usually increase the excitation energies computed at non-relativistic level, leading to a better agreement with the experimental values.<sup>104</sup> SO coupling effects arise from the interaction between the electronic magnetic moment and the magnetic field generated by its own orbital motion; their inclusion leads to doubly-split states deriving from excitations originating from a non-*s* ( $l \neq 0$ ) core orbital. Therefore, the electronic structure is described in terms of spinors (*i.e.*, one-electron functions expressed as sum of two products between a spatial and a spin function); this causes a reduction of the degeneracy going from MOs to SO spinors. Depending on the SO value, the analysis of computed core excitation spectra can be more or less demanding: for instance, in the case of 2p transitions, as those investigated in this thesis work, the two manifolds of excited states converging to the L<sub>III</sub> (2p<sub>3/2</sub>) and L<sub>II</sub> (2p<sub>1/2</sub>) edges strongly overlap, rendering the assignment of the calculated peaks quite demanding. An additional complication derives from the molecular-field splitting originated from the 2p<sub>3/2</sub> core level. It is therefore evident that both the coupling between different excitation channels arising from the 2p degenerate core-holes and relativistic SO coupling effects are essential to obtain a quantitative description of the series of transitions converging to the L<sub>II</sub> and L<sub>III</sub>-edges. In this thesis work,

TDDFT and the relativistic two-component *Zeroth-Order Regular Approximation (ZORA)*<sup>105-107</sup> have been employed to simulate the NEXAFS spectra at the  $L_{II,III}$ -edge region.

The following paragraphs provide a derivation of the CASIDA formulation of TDDFT in the linear response regime<sup>26</sup>, starting from the bases of Linear Response Theory<sup>108</sup>. An overview of the main steps of the TDDFT formalism when dealing with core electron excitations will be finally given.

### 2.7.1 Response Theory (RT)

Response Theory aims at describing the way in which a system reacts to external perturbations, such as applied electric or magnetic fields. Consider a  $N$ -electron system which satisfies the time-dependent Schrödinger equation:

$$H\Psi = i \frac{\partial \Psi}{\partial t} \quad (2.97)$$

If the time-dependent perturbation is small the perturbed Hamiltonian assumes the following form:

$$H(t) = H^{(0)} + \lambda H^{(1)}(t) \quad (2.98)$$

where the superscripts indicate the approximation order, and  $\lambda$  denotes the parameter that “switches on” the perturbation. The perturbed wave function is also expanded in powers of  $\lambda$ :

$$\Psi'(t) = \Psi^{(0)} + \lambda \Psi^{(1)}(t). \quad (2.99)$$

In this treatment, we assume that the unperturbed problem has already been solved, *i.e.*, that we can calculate the eigenvectors and eigenvalues through solution of the time-independent Schrödinger equation. Since in general the unperturbed wave functions  $\Psi_n^{(0)}$  form a complete set, it is possible to expand the perturbed wave function, for  $t > 0$ , in terms of them:

$$\Psi'(\mathbf{r}, t) = \sum_{n \neq 0} c_n(t) \Psi_n^{(0)}(\mathbf{r}) e^{-iE_n t}, \quad (2.100)$$

where the  $c_n(t)$ 's are the expansion coefficients that contain the evolution in time of the wave function, and  $e^{-iE_n t}$  is an exponential phase factor.

The unperturbed wave function can be factorized in two parts, namely, one spatial and one depending on time, as follows:

$$\Psi_n^{(0)}(\mathbf{r}, t) = \Psi_n^{(0)}(\mathbf{r}) e^{-iE_n t}. \quad (2.101)$$

Therefore, the solution of the perturbed system is the following one:

$$|\Psi'_0\rangle = \Psi'_0(\mathbf{r}, t)e^{-iE_0t} = \Psi_0^{(0)}(\mathbf{r})e^{-iE_0t} + \lambda \sum_{n \neq 0} c_n(t) \Psi_n^{(0)}(\mathbf{r})e^{-iE_n t}. \quad (2.102)$$

If  $\omega_{0n} = E_n - E_0$ , then

$$\Psi'_0(\mathbf{r}, t) = \Psi_0^{(0)}(\mathbf{r}) + \lambda \sum_{n \neq 0} c_n(t) \Psi_n^{(0)}(\mathbf{r})e^{-i\omega_{0n}t}. \quad (2.103)$$

Introducing Eq. (2.103) in the time-dependent Schrödinger equation (Eq. 2.97) and considering the first order in the expansion series, if one collects the exponential terms:

$$H^{(1)} \Psi_0^{(0)} = i \sum_{n \neq 0} \dot{c}_n \Psi_n^{(0)} e^{-i\omega_{0n}t}. \quad (2.104)$$

If one multiplies Eq. (2.104) for the bra element  $\langle \Psi_k^{(0)} |$ , the coefficients  $c_k$  can be found through the resolution of the following integral:

$$c_k(t) = \int_{-\infty}^t \dot{c}_k dt = -i \int_{-\infty}^t \langle \Psi_k^{(0)} | H^{(1)} | \Psi_0^{(0)} \rangle e^{i\omega_{0k}t'} dt', \quad (2.105)$$

where we write the perturbation term  $H^{(1)}$  as  $H^{(1)}(t) = A(\mathbf{r})F(t)$ , with  $A$  only depending on space, and  $F$  only on time. Therefore, the integral can be split in two parts, namely, a spatial one, from which the shape of the perturbation can be extracted, and one depending on time and frequency, as follows:

$$c_f(t) = -i \langle n | A | 0 \rangle \int_{-\infty}^t F(t') e^{i\omega_{0n}t'} dt'. \quad (2.106)$$

The coefficients include the contribution of the excited state, due to the perturbation, on the first order perturbed ground state wave function.

It is usually more useful to study the effect of the perturbation on specific properties of the wave function, *i.e.*, the actual response of the system. The latter can be evaluated through the variation of the expectation value  $\langle B \rangle$  of the observable  $B$ , generated by a perturbation  $A$ , between the perturbed function and the initial one; the first order term is the linear response (*i.e.*, linear in  $F(t)$ ).

$$\delta \langle B(t) \rangle = \langle B(t) \rangle - \langle B \rangle_0. \quad (2.107)$$

By employing Eq. (2.104), one obtains:

$$\delta \langle B(t) \rangle = \langle \Psi | B | \Psi \rangle - \langle \Psi^{(0)} | B | \Psi^{(0)} \rangle. \quad (2.108)$$

By using Eq. (2.99), Eq. (2.108) becomes (at first order in the perturbation):

$$\begin{aligned}
\delta\langle B(t)\rangle &= \langle \Psi^{(0)} + \lambda\Psi^{(1)}|B|\Psi^{(0)} + \lambda\Psi^{(1)}\rangle - \langle \Psi^{(0)}|B|\Psi^{(0)}\rangle = \\
&= \langle \Psi^{(0)}|B|\Psi^{(0)}\rangle + \langle \Psi^{(0)}|B|\lambda\Psi^{(1)}\rangle + \langle \lambda\Psi^{(1)}|B|\Psi^{(0)}\rangle + \langle \lambda\Psi^{(1)}|B|\lambda\Psi^{(1)}\rangle - \langle \Psi^{(0)}|B|\Psi^{(0)}\rangle = \\
&= \langle \Psi^{(0)}|B|\lambda\Psi^{(1)}\rangle + \langle \lambda\Psi^{(1)}|B|\Psi^{(0)}\rangle.
\end{aligned} \tag{2.109}$$

By replacing the perturbed eigenfunction with its expansion in terms of unperturbed wave functions, one obtains:

$$\delta\langle B(t)\rangle = \sum_{n\neq 0} c_n(t)\langle 0|B|n\rangle e^{-i\omega_{0n}t} + c.c. \tag{2.110}$$

If the expression for the coefficients in Eq. (2.106) is used, Eq. (2.110) becomes:

$$\delta\langle B(t)\rangle = -i \sum_{n\neq 0} \langle 0|B|n\rangle \langle n|A|0\rangle \int_{-\infty}^t F(t') e^{-i\omega_{0n}(t-t')} dt' + c.c., \tag{2.111}$$

which can be simplified by introducing the “time-correlation function”  $K$ :

$$\delta\langle B(t)\rangle = \int_{-\infty}^t K(BA|t-t') F(t') dt' + c.c., \tag{2.112}$$

with  $K(BA|t-t') = -i \sum_{n\neq 0} \langle 0|B|n\rangle \langle n|A|0\rangle e^{-i\omega_{0n}(t-t')}$ . The perturbation  $A$  at time  $t'$  is related to the oscillations of the observable  $B$  at time  $t$ ; according to the causality principle,  $t > t'$ , otherwise  $K(BA|t-t')$  is not defined.

In Linear Response Theory, it is convenient to work in the frequency domain. Hence, it is useful to consider the frequency components of  $F(t)$  obtained through a Fourier transform (FT)  $f(\omega)$ . Before operating through the FT, it is useful to find an expression for  $F(t)$  so that the perturbation vanishes in remote times (this ensures to start from unperturbed states); to do this, one introduces an adiabatic “switch on” factor  $e^{\varepsilon t}$  such that:

$$\lim_{\substack{\varepsilon \rightarrow 0^+ \\ t \rightarrow -\infty}} e^{\varepsilon t} = 0 \tag{2.113}$$

while for finite  $t$ ,  $e^{\varepsilon t} \approx 1$ .

The  $\varepsilon \rightarrow 0^+$  condition ensures a gradual “switch on” of the perturbation to avoid transient effects.



The FT of  $F(t)$  is given by:

$$F(t) = e^{\varepsilon t} \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) e^{-i\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) e^{-i(\omega+i\varepsilon)t} d\omega. \quad (2.114)$$

By replacing Eq. (2.115) in Eq. (2.112), one obtains:

$$\begin{aligned} \delta\langle B(t) \rangle = & -\frac{i}{2\pi} \int_{-\infty}^t \int_{-\infty}^{+\infty} f(\omega) e^{-i(\omega+i\varepsilon)t'} \sum_{n \neq 0} [\langle 0|B|n\rangle \langle n|A|0\rangle e^{-i\omega_{0n}(t-t')} \\ & - \langle n|B|0\rangle \langle 0|A|n\rangle e^{i\omega_{0n}(t-t')}] dt' d\omega \end{aligned} \quad (2.115)$$

The integral in time involves only the exponential factors, therefore:

$$\int_{-\infty}^t e^{-i(\omega+i\varepsilon-\omega_{0n})t'} e^{-i\omega_{0n}t} dt' = e^{i\omega_{0n}t} \frac{e^{-i(\omega+i\varepsilon-\omega_{0n})t'}}{-i(\omega+i\varepsilon-\omega_{0n})} \Big|_{-\infty}^t = -\frac{e^{-i(\omega+i\varepsilon)t}}{i(\omega+i\varepsilon-\omega_{0n})}$$

and

$$\int_{-\infty}^t e^{-i(\omega+i\varepsilon+\omega_{0n})t'} e^{i\omega_{0n}t} dt' = e^{i\omega_{0n}t} \frac{e^{-i(\omega+i\varepsilon+\omega_{0n})t'}}{-i(\omega+i\varepsilon+\omega_{0n})} \Big|_{-\infty}^t = -\frac{e^{-i(\omega+i\varepsilon)t}}{i(\omega+i\varepsilon+\omega_{0n})} \quad (2.116)$$

By replacing Eq. (2.116) in Eq. (2.115), one obtains:

$$\delta\langle B(t) \rangle = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) e^{-i(\omega+i\varepsilon)t} \sum_{n \neq 0} \left[ \frac{\langle 0|B|n\rangle \langle n|A|0\rangle}{\omega+i\varepsilon-\omega_{0n}} - \frac{\langle n|B|0\rangle \langle 0|A|n\rangle}{\omega+i\varepsilon+\omega_{0n}} \right] d\omega, \quad (2.117)$$

where the sum is the FT of the time-correlation function,  $K$ , and defines the dynamic polarizability:

$$\Pi(BA|\omega) = \sum_{n \neq 0} \left[ \frac{\langle 0|B|n\rangle \langle n|A|0\rangle}{\omega+i\varepsilon-\omega_{0n}} - \frac{\langle n|B|0\rangle \langle 0|A|n\rangle}{\omega+i\varepsilon+\omega_{0n}} \right] \quad (2.118)$$

Therefore, Eq. (2.117) can be rewritten as:

$$\delta\langle B(t) \rangle = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) e^{-i(\omega+i\varepsilon)t} \Pi(BA|\omega) d\omega. \quad (2.119)$$

It is interesting to note that the dynamic polarizability has two poles in  $\omega = \omega_{0n} - i\varepsilon$  and  $\omega = -i\varepsilon - \omega_{0n}$ . In resonant conditions (*i.e.*,  $\omega = \omega_{0n}$ ):

$$\lim_{\varepsilon \rightarrow 0^+} \Pi(BA|\omega_{0n}) \cong \frac{\langle 0|B|n\rangle \langle n|A|0\rangle}{i\varepsilon}. \quad (2.120)$$

Therefore, in resonance conditions, the dynamic polarizability is an imaginary number. If one considers the electronic dipoles induced by the oscillating electric field of the electromagnetic radiation, one obtains:

$$Im[\lim_{\varepsilon \rightarrow 0^+} (\mu\mu | \omega_{0n})] \cong -\frac{|\langle n|\mu|0\rangle|^2}{\varepsilon} \quad (2.121)$$

Hence, the polarizability, denoted by  $\alpha$ , is proportional to the transition moment. The problem can be then solved in two different ways. The first one involves the resolution of the response matrix by means of a diagonalization, thus obtaining a discrete spectrum, while in the second one, the spectrum is calculated point by point from the imaginary part of the polarizability, setting  $\varepsilon$  as a finite number.

Let's consider a monochromatic perturbation of  $\omega_0$  pulsation induced by an electromagnetic radiation:

$$V(\mathbf{r}, t) = A(\mathbf{r})F(t) = A(\mathbf{r})\cos(\omega_0 t) \quad (2.122)$$

By applying the FT (Eq. 2.114), one obtains:

$$\begin{aligned} f(\omega) &= \int_{-\infty}^{+\infty} F(t)e^{-i\omega t} dt = \int_{-\infty}^{+\infty} \left( \frac{e^{i\omega_0 t} + e^{-i\omega_0 t}}{2} \right) e^{-i\omega t} dt = \\ &= \frac{1}{2} \left[ \int_{-\infty}^{+\infty} e^{-i(\omega + \omega_0)t} dt + \int_{-\infty}^{+\infty} e^{-i(\omega - \omega_0)t} dt \right] = \\ &= \frac{1}{2} [2\pi\delta(\omega + \omega_0) + 2\pi\delta(\omega - \omega_0)] = \pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] \end{aligned} \quad (2.123)$$

Therefore, the effect of the FT on  $F(t)$  is:

$$F(t) = \cos(\omega_0 t) \xrightarrow{FT} \pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]. \quad (2.124)$$

Now, it is useful to see the effect of the Dirac delta function, by verifying that the expression found for the FT of  $F(t)$  is correct; this can be done by considering the Fourier anti-transform:

$$\begin{aligned} F(t) = \cos(\omega_0 t) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) e^{i\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] e^{i\omega t} d\omega = \\ &= \frac{1}{2} \int_{-\infty}^{+\infty} \delta(\omega - \omega_0) e^{i\omega t} d\omega + \frac{1}{2} \int_{-\infty}^{+\infty} \delta(\omega + \omega_0) e^{i\omega t} d\omega = \frac{e^{-i\omega_0 t} + e^{i\omega_0 t}}{2} \end{aligned} \quad (2.125)$$

By using the expression found for  $f(\omega)$  (Eq. 2.123) in Eq. (2.119); one obtains:

$$\begin{aligned}
\delta\langle B(t)\rangle &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)] e^{-i(\omega+i\varepsilon)t} \alpha(BA|\omega) d\omega = \\
&= \frac{1}{2} e^{-i\omega_0 t} \sum_{n \neq 0} \left[ \frac{\langle 0|B|n\rangle\langle n|A|0\rangle}{\omega_0 - \omega_{0n}} - \frac{\langle n|B|0\rangle\langle 0|A|n\rangle}{\omega_0 + \omega_{0n}} \right] + \frac{1}{2} e^{i\omega_0 t} \sum_{n \neq 0} \left[ \frac{\langle 0|B|n\rangle\langle n|A|0\rangle}{-\omega_0 - \omega_{0n}} - \frac{\langle n|B|0\rangle\langle 0|A|n\rangle}{-\omega_0 + \omega_{0n}} \right] = \\
&= \frac{1}{2} \left[ e^{-i\omega_0 t} \Pi(BA|\omega_0) + e^{i\omega_0 t} \Pi(BA|-\omega_0) \right]
\end{aligned} \tag{2.126}$$

The first order perturbed Hamiltonian can be expressed as:

$$H'(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\omega) \frac{1}{2} [A_\omega e^{-i\omega t} - A_{-\omega} e^{i\omega t}] d\omega. \tag{2.127}$$

In the case of a monochromatic perturbation with  $\omega$  frequency:

$$H'(\omega) = \frac{1}{2} [A_\omega e^{-i\omega t} - A_{-\omega} e^{i\omega t}]. \tag{2.128}$$

The coefficients of the perturbed wave function can be expressed in frequency terms by means of the Fourier anti-transform:

$$c_n(t) = -\frac{i}{2} \int_{-\infty}^t \langle n|A_\omega|0\rangle e^{i(\omega_{0n} - \omega - i\varepsilon)t'} dt' + \text{term with } \omega \rightarrow -\omega \tag{2.129}$$

whose integration gives:

$$c_n(t) = -\frac{1}{2} \left[ \langle n|A_\omega|0\rangle \frac{e^{i(\omega_{0n} - \omega - i\varepsilon)t}}{\omega_{0n} - \omega - i\varepsilon} + \text{term with } \omega \rightarrow -\omega \right]. \tag{2.130}$$

Since these expressions can be obtained only in the limit  $\varepsilon \rightarrow 0^+$ , the “switch on” factor  $\varepsilon^{et}$  is important only at the denominator, and can be interpreted as the imaginary part of the photon frequency. Moreover, in resonance conditions (*i.e.*,  $\omega = \omega_{0n}$ ), the expression in Eq. (2.130) tends to infinity, and a discrete line spectrum is obtained. If a small finite  $\varepsilon$  is employed, one gets a spectrum convoluted with Lorentzian functions.

By using Eq. (2.126) and expressing the oscillations of  $\langle B \rangle$  as a function of the frequency oscillations, one obtains:

$$\delta\langle B(\omega)\rangle = \frac{1}{2} \left[ e^{-i\omega t} \Pi(BA_\omega|\omega) + e^{i\omega t} \Pi(BA_\omega|-\omega) \right] \tag{2.131}$$

It is worth noting that the expressions obtained in this treatment cannot be however applied in practice, since the involved infinite summations on all the exact excited states are not accessible in practice. By applying this formalism on simplified, model descriptions of the system, such as Hartree-Fock or Kohn-Sham, it is possible to obtain a method to be implemented and employed in practice; this will be the subject of the next paragraph

## 2.7.2 Random Phase Approximation (RPA) and TDDFT

With reference to the first order Time Dependent-Perturbation Theory (TD-PT) (Eq. 2.127) and considering a perturbing term with  $\omega$  frequency, one obtains:

$$H^{(1)}(t) = A_{\omega}e^{-i\omega t} + A_{-\omega}e^{i\omega t} \quad (2.132)$$

Assuming that the first order perturbed wave function has only two different “time behaviors”, namely,  $e^{-i(E_0+\omega)t}$  and  $e^{-i(E_0-\omega)t}$ , due to the two different terms which appear in  $H^{(1)}$ , the time-dependent Schrödinger equation, Eq. (2.97), becomes:

$$\begin{aligned} [H^{(0)} + \lambda(A_{\omega}e^{-i\omega t} + A_{-\omega}e^{i\omega t})][\Psi_0^0(\mathbf{r})e^{-iE_0t} + \lambda(\Psi^{(1,-)}(\mathbf{r})e^{-i(E_0+\omega)t} + \Psi^{(1,+)}(\mathbf{r})e^{-i(E_0-\omega)t})] = \\ = i \frac{\partial}{\partial t} [\Psi_0^0(\mathbf{r})e^{-iE_0t} + \lambda(\Psi^{(1,-)}(\mathbf{r})e^{-i(E_0+\omega)t} + \Psi^{(1,+)}(\mathbf{r})e^{-i(E_0-\omega)t})] \end{aligned} \quad (2.133)$$

where  $\Psi^{(1,-)}$  and  $\Psi^{(1,+)}$  denote the first order perturbed wave function at the two different “time behaviors”.

The zero order gives the time-independent equation:

$$H^{(0)} \Psi_0^{(0)}(\mathbf{r})e^{-iE_0t} = E_0 \Psi_0^{(0)}(\mathbf{r})e^{-iE_0t}, \quad (2.134)$$

while the first order terms can be grouped in two linear independent, time dependent contributions, namely, one for  $e^{-i(E_0+\omega)t}$ , and one for  $e^{-i(E_0-\omega)t}$ , respectively:

$$\begin{aligned} e^{-i(E_0+\omega)t} [H^{(0)}\Psi^{(1,-)}(\mathbf{r}) + A_{\omega} \Psi_0^{(0)}(\mathbf{r}) - (E_0 + \omega)\Psi^{(1,-)}(\mathbf{r})] + \\ e^{-i(E_0-\omega)t} [H^{(0)}\Psi^{(1,+)}(\mathbf{r}) + A_{-\omega} \Psi_0^{(0)}(\mathbf{r}) - (E_0 - \omega)\Psi^{(1,+)}(\mathbf{r})] = 0 \end{aligned} \quad (2.135)$$

Both expressions in brackets vanish, since the whole expression must be identically zero.

Let's consider now the special case of Hartree-Fock, where  $H^{(0)} \equiv \hat{F}$ . The perturbed wave functions must preserve the orthonormality condition,  $\langle \varphi_i | \varphi_j \rangle = \delta_{ij}$ , where indexes  $i, j$  denote the occupied orbitals. Hence:

$$\langle \delta \varphi_i | \varphi_j \rangle + \langle \varphi_i | \delta \varphi_j \rangle = 0. \quad (2.136)$$

Therefore, each term is equal to zero; this implies that the perturbed function is orthogonal to all the occupied orbitals, and hence must be a linear combination of virtual orbitals:

$$|\delta \varphi_i\rangle = \sum_b^{virt} c_{ib} |\varphi_b\rangle. \quad (2.137)$$

This latter expression must satisfy both the “time behaviors” of the perturbed wave function, which can be expressed in the following way:

$$\Psi_i^{(1,-)} = \sum_b^{virt} X_{ib} \varphi_b \quad \Psi_i^{(1,+)} = \sum_b^{virt} Y_{ib}^* \varphi_b \quad (2.138)$$

Taking back the general equations of the first order Time Dependent-Dynamic Perturbation (TD-DP) theory, and doing a scalar product with a bra  $\langle \varphi_a |$ , where the  $a$  index denotes a virtual orbital, one gets the following expressions:

$$\sum_b \langle \varphi_a | F | \varphi_b \rangle X_{ib} + \langle \varphi_a | A_\omega | \varphi_i \rangle - (\varepsilon_i + \omega) \sum_b \langle \varphi_a | \varphi_b \rangle X_{ib} = 0 \quad (2.139)$$

$$\sum_b \langle \varphi_a | F | \varphi_b \rangle Y_{ib}^* + \langle \varphi_a | A_{-\omega} | \varphi_i \rangle - (\varepsilon_i - \omega) \sum_b \langle \varphi_a | \varphi_b \rangle Y_{ib}^* = 0 \quad (2.140)$$

where Eq. (2.139) corresponds to the first term of Eq. (2.135), while Eq. (2.140) corresponds to the second term of Eq. (2.135).

In both expressions, the sum reduces to one term only, which correspond to  $b = a$ . Transforming Eq. (2.140) in its conjugated complex, other two simpler expressions are obtained:

$$(\varepsilon_a - \varepsilon_i - \omega) X_{ia} + \langle \varphi_a | A_\omega | \varphi_i \rangle = 0 \quad (2.141)$$

$$(\varepsilon_a - \varepsilon_i + \omega) Y_{ia} + \langle \varphi_i | A_\omega | \varphi_a \rangle = 0 \quad (2.142)$$

The TD external field, whose strength corresponds to the  $A_\omega$  operator, affects the orbitals; as a consequence, also the Fock operator changes. The  $\langle \varphi_a | A_\omega | \varphi_i \rangle$  term can be evaluated by considering that  $\langle \varphi_a | F^{(1)} | \varphi_i \rangle = \langle a | F^{(1)} | i \rangle = \sum_j (\langle a \delta_j | | i j \rangle + \langle a j | | i \delta_j \rangle)$ , where  $F^{(1)} = \sum_j^{occ} (\langle \delta_j | | j \rangle + \langle j | | \delta_j \rangle)$  and both  $j$  and  $\delta j$  are expressed with their TD components. Therefore,

$$\langle a | F^{(1)} | i \rangle = \sum_{jb} e^{-i\omega t} [Y_{jb} \langle ab | | i j \rangle + X_{jb} \langle a j | | i b \rangle] + \sum_{jb} e^{i\omega t} [X_{ib}^* \langle ab | | i j \rangle + Y_{jb}^* \langle a j | | i b \rangle]. \quad (2.143)$$

Hence, the  $\langle \varphi_a | A_\omega | \varphi_i \rangle$  term is equal to:

$$\langle \varphi_a | A_\omega | \varphi_i \rangle = \langle a | z | i \rangle + \sum_{jb} \langle a j | | i b \rangle X_{jb} + \langle a b | | i j \rangle Y_{jb} = -\langle a | z | i \rangle \quad (2.144)$$

where the  $X$  and  $Y$  coefficients are unknown.

Equations (2.141) and (2.142) can be therefore rewritten as follows:

$$(\varepsilon_a - \varepsilon_i - \omega) X_{ia} + \sum_{jb} \langle a j | | i b \rangle X_{jb} + \langle a b | | i j \rangle Y_{jb} = -\langle a | z | i \rangle \quad (2.145)$$

$$(\varepsilon_a - \varepsilon_i + \omega) Y_{ia} + \sum_{jb} \langle i b | | a j \rangle Y_{jb} + \langle i j | | a b \rangle X_{jb} = -\langle a | z | i \rangle \quad (2.146)$$

and both represented in matrix form:

$$\begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} - \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} -V \\ -V \end{pmatrix} \quad (2.147)$$

where  $A_{ia,jb} = \delta_{ij} \delta_{ab} (\varepsilon_a - \varepsilon_i) + \langle a j | | i b \rangle$  and  $B_{ia,jb} = \langle a b | | i j \rangle$ .

From equations (2.145) and (2.146) the coefficients  $X$  and  $Y$  for a given frequency  $\omega$  can be extracted; when the frequency corresponds to an excitation energy, the polarizability goes to infinity.

To determine the response of the system, it is necessary to consider the expression of  $\delta\langle B(t) \rangle$ , Eq. (2.107). One obtains, for  $B = z$ , the dipole component along  $z$ :

$$\delta\langle B(t) \rangle = \sum_i (\langle \delta\varphi_i | \hat{B} | \varphi_i \rangle + \langle \varphi_i | \hat{B} | \delta\varphi_i \rangle) = (\langle \delta\varphi_i | z | \varphi_i \rangle + \langle \varphi_i | z | \delta\varphi_i \rangle). \quad (2.148)$$

Considering that the first-order perturbation of the orbitals,  $\delta\varphi_i$ , can be expressed as a linear combination of virtual orbitals (Eq. 2.137), Eq. (2.148) can be rewritten as:

$$\delta\langle B(t) \rangle = \sum_{i,a} e^{-i\omega t} [Y_{ia} \langle a | z | i \rangle Y_{ia} + X_{ia} \langle i | z | a \rangle] + \sum_{i,a} e^{i\omega t} [X_{ia}^* \langle a | z | i \rangle Y_{ia} + Y_{ia}^* \langle i | z | a \rangle]. \quad (2.149)$$

By comparison with the expression of  $\delta\langle B_\omega \rangle$  (Eq. 2.131), it is possible to identify:

$$\Pi(BA_\omega | \omega) = \sum_{i,a} Y_{ia} \langle a | z | i \rangle + X_{ia} \langle i | z | a \rangle. \quad (2.150)$$

$\delta\langle B(t) \rangle$  (Eq. 2.148) has poles when the  $X$  and  $Y$  coefficients are the solutions of the pseudo-eigenvalue equations:

$$\begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \quad (2.151)$$

The expression above corresponds to the *Random Phase Approximation* (RPA) equation.

The problem can be simplified in two cases, namely, when orbitals are real (and therefore  $A = A^*$  and  $B = B^*$ ), and when the  $B$  matrix can be neglected (the so-called *Tamm Dancoff Approximation-TDA*), which leads to the eigenvalue equation  $AX = \omega X$ .

If the Kohn-Sham operator  $\hat{H}_{KS}$  is used as  $H^{(0)}$  instead of the Fock operator  $\hat{F}$ , starting from Eq. (2.147), it is possible to switch from RPA to TDDFT. The matrix elements of  $A$  and  $B$  differ from the RPA ones only for the presence of bi-electronic term, since in the KS approach the exchange-correlation potential  $v_{xc}$  must be included. In particular, the elements of  $A$  and  $B$  matrices for RPA are the following ones:

$$A_{ia,jb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + \langle aj||ib \rangle = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + \langle aj|ib \rangle - \langle aj|bi \rangle, \quad (2.152)$$

and

$$B_{ia,jb} = \langle ab||ij \rangle = \langle ab|ij \rangle - \langle ab|ji \rangle, \quad (2.153)$$

while for TDDFT, they assume the following form:

$$A_{ia,jb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + \langle aj||ib \rangle = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i) + \langle aj|ib \rangle - \langle aj|K_{xc}|bi \rangle \quad (2.154)$$

$$B_{ia,jb} = \langle ab||ij \rangle = \langle ab|ij \rangle - \langle ab|K_{xc}|ij \rangle. \quad (2.155)$$

If both  $A$  and  $B$  are real matrices (which is the case for real MOs), then  $\langle aj|ib \rangle = \langle ab||ij \rangle$ . Eq. (2.147) is therefore simplified to a system of equations:

$$\begin{cases} AX + BY - \omega X = -V \\ BX + AY + \omega X = -V \end{cases} \quad (2.156)$$

Finding the value of  $(X - Y) = \omega(A - B)^{-1}(X + Y)$ , the following equation is found:

$$(A + B) - \omega^2(A - B)^{-1}(X + Y) = -2V \quad (2.157)$$

where  $(A - B)^{-1}$  is diagonal in TDDFT, and can be redefined as  $\boldsymbol{\varepsilon}^{-1}$ , where  $\varepsilon_{ia,jb} = \delta_{ij}\delta_{ab}(\varepsilon_a - \varepsilon_i)$ .

The final expression leads to the eigenvalue equation in the Casida formulation of TDDFT<sup>26</sup>:

$$\Omega F = \omega^2 F \quad (2.158)$$

where  $\Omega \equiv \varepsilon^{-\frac{1}{2}}(A + B)\varepsilon^{\frac{1}{2}}$  and  $F \equiv \varepsilon^{-\frac{1}{2}}(X + Y)$ .

The problem is therefore reduced to an eigenvalue equation that gives  $(\Omega - \omega^2)^{-1} = F_p \left( \frac{1}{\omega_p^2 - \omega^2} \right) F_p^\dagger$ .

The result of this perturbation on a property of the system is the following one:

$$\delta\langle B \rangle = V^\dagger (X - Y) = V^\dagger \varepsilon^{\frac{1}{2}} F = V^\dagger \varepsilon^{-\frac{1}{2}} (-2) F_p \left( \frac{1}{\omega_p^2 - \omega^2} \right) F_p^\dagger \varepsilon^{\frac{1}{2}} V \quad (2.159)$$

When the frequency of the external potential is resonant with one of the excited state energy ( $\omega \rightarrow \omega_p$ ), the oscillations of the expectation value  $\langle B \rangle$  lead to the following expression:

$$\delta\langle B \rangle \cong \left| V^\dagger \varepsilon^{\frac{1}{2}} F_p \right|^2 \frac{(-2)}{\omega_p^2 - \omega^2} \cong \frac{\left| V^\dagger \varepsilon^{\frac{1}{2}} F_p \right|^2 (-2)}{(\omega_p - \omega) 2\omega} \cong \frac{\left| V^\dagger \varepsilon^{\frac{1}{2}} F_p \right|^2}{(\omega - \omega_p) \omega} = \frac{|\langle p|z|0 \rangle|^2}{\omega - \omega_p} \quad (2.160)$$

where  $\omega |\langle p|z|0 \rangle|^2 = \left| V^\dagger \varepsilon^{\frac{1}{2}} F_p \right|^2$

Hence, in the Casida formulation of TDDFT, the solutions of the problem are found through the diagonalization of the matrix  $\Omega$ , which is a four-indexed matrix defined within the 1h-1p space. The dimension of  $\Omega$  is given by the product of the number of occupied and virtual orbitals, while its elements are:

$$\Omega_{ia\sigma, bj\tau} = \delta_{\sigma\tau} \delta_{ij} \delta_{ab} (\varepsilon_a - \varepsilon_i)^2 + 2\sqrt{(\varepsilon_a - \varepsilon_i)} \frac{\partial F_{ia}}{\partial P_{jb}} \sqrt{(\varepsilon_b - \varepsilon_j)} \quad (2.161)$$

where indices  $i$  and  $j$  run over the set of occupied spinors in the KS ground state, while indices  $a$  and  $b$  run over the set of virtual spinors.  $\sigma$  and  $\tau$  are the indices for the spin, and  $\varepsilon_i$  and  $\varepsilon_a$  are the KS molecular orbital energies. F and P represent the Fock matrix and the density matrix, respectively, whereas  $\frac{\partial F_{ia}}{\partial P_{jb}}$  are the elements of the coupling matrix K:

$$K_{ia\sigma, bj\tau} = \int d\mathbf{r} \int d\mathbf{r}' \varphi_{i\sigma}(\mathbf{r}) \varphi_{a\sigma}(\mathbf{r}) \left[ \frac{1}{|\mathbf{r} - \mathbf{r}'|} + f_{xc}^{ALDA}(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}') \right] \varphi_{j\tau}(\mathbf{r}') \varphi_{b\tau}(\mathbf{r}') \quad (2.162)$$

with  $f_{xc}^{ALDA}(\mathbf{r})$  is the exchange-correlation kernel approximated by using the *Adiabatic Local Density Approximation* (ALDA, [109]), which represents the simplest approximation in TDDFT to treat  $f_{xc}$ .

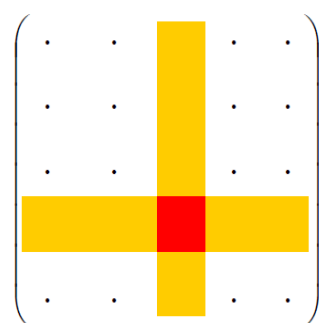
The intensity of the spectral lines, corresponding to the excitation energies given by  $\omega_p$ , is expressed by the oscillator strengths that are extracted from the eigenvectors  $F_p$ :

$$F_{pI} \omega = \frac{2}{3} (|\langle p|x|0 \rangle|^2 + |\langle p|y|0 \rangle|^2 + |\langle p|z|0 \rangle|^2) = \frac{2}{3} \left( x^\dagger \varepsilon^{\frac{1}{2}} F_p + y^\dagger \varepsilon^{\frac{1}{2}} F_p + z^\dagger \varepsilon^{\frac{1}{2}} F_p \right). \quad (2.163)$$

However, the direct solution of the eigenvalue equation (Eq. 2.158) is possible exclusively in principle, but it is infeasible due to computational and storage requirements; therefore, in the case of large molecules, it is preferable to solve the eigenvalue problem iteratively by employing the



Davidson's algorithm<sup>110</sup>, which requires low computational costs also for large matrices. This method is however limited to a few number of selected eigenvalues, corresponding to the lowest excitation energies; therefore, some approximations must be adopted to obtain core excitation energies, which lie very high in the excitation spectrum. In this respect, the core–valence separation approximation (CVS) by Cederbaum *et al.*<sup>111,112</sup> can be employed; within this approximation, one can decouple core orbitals from the valence orbitals, so that the respective blocks of the model Hamiltonian are zeroed out (*i.e.*, the solutions of the Schrödinger equation can be separated into the core and valence domains). This is possible due to the large energy separation between the valence and core excited states.<sup>113</sup> The mono-excited (1h-1p) configuration space can be hence described by occupied-virtual couples of indexes (*e.g.*,  $ia$  or  $jb$ ); since the occupied index  $i$  is allowed to span only the core orbitals, and no limitations are set for the virtual index  $a$ , all and only core excitations are included. In the case of 2p excitations, like those treated in this thesis work, the indices of occupied spinors run over the  $2p_{1/2}$  and  $2p_{3/2}$  subshells, hence, allowing a mixing of only those initial states; this leads to a consequent drastic reduction of the  $\Omega$  matrix dimensions. The eigenvalue equation can be therefore efficiently solved with Davidson's iterative algorithm, since the core excitations now correspond to the lowest roots of the  $\Omega$  matrix; this also results in a considerable computational savings (see Figure 2.2).



**Figure 2.2** – The core-valence separation allows a considerable reduction of the dimensions of the  $\Omega$  matrix.

Finally, it is necessary to point out that the calculation of NEXAFS spectra through TDDFT is strongly affected by the maximum number of roots of Eq. (2.158) that can be safely computed iteratively. Indeed, in the case of large molecular systems, the core-excited states are so close to each other in energy that the number of roots necessary to cover the excitation energy range for a complete simulation of the experiment is considerably high ( $> 1000$ ); in such cases, the Davidson's iterative algorithm could become numerically unstable, even within the CVS approximation.

## 2.8 Theoretical methods for core-electron excitation computations

In this thesis work, DFT within the Transition Potential (TP) scheme has been employed for the computation of *K-edge* NEXAFS spectra, and its time-dependent generalization (TDDFT) in the linear response regime in the case of *L<sub>II,III</sub>-edge* NEXAFS spectra to include both spin-orbit coupling effects and configuration mixing among different excitation channels from the degenerate 2p core-holes. These choices are justified by considering that the accuracy of the DFT-TP method in the half core hole approximation (HCH) for *K-edge* NEXAFS spectra of first-row atoms is usually comparable (if not greater) than that afforded by Configuration Interaction Singles (CIS, [114])/TDDFT methods<sup>115-117</sup>. The reasons behind the effectiveness of single particle approaches in the description of *K-edge* NEXAFS spectra can be identified in the weak coupling of core excitations from non-degenerate core-holes as well as in the importance of relaxation effects which are explicitly taken into account in the HCH approximation, but cannot be always accurately described by including only the manifold of single excitations from the Hartree-Fock (HF)/Kohn-Sham (KS) Ground State determinant<sup>118</sup>. Furthermore, it is well known that the configuration mixing among different excitation channels from degenerate core-holes and spin-orbit coupling effects must be considered to provide a reliable description of *L<sub>II,III</sub>-edge* NEXAFS spectra of heavy atoms.<sup>119-121</sup>

Density Functional Theory (DFT) is the most robust and popular method employed for core-electron excitation calculations. Computational approaches involving pseudo-potentials are also well exploited, as they allow the inclusion of relativistic effects in DFT methods in a more efficient way compared to that provided by all-electron methods. Furthermore, the pseudopotential approximation significantly lowers the computational effort in the calculations of large molecules with a considerable reduction of the basis set, since integrals deriving from core orbitals are substituted by corresponding (one-electron) pseudopotential integrals<sup>122</sup>. A powerful strategy to calculate the electronic structure of molecules and relatively large clusters (up to 50 atoms) is represented by the self-consistent field X $\alpha$ -Multiple Scattering (MS) method, devised by Slater<sup>25</sup> and Johnson<sup>123</sup> in the late 1960s, and originally employed for bound electronic states. In the middle 1970s, this method was extended to the description of continuum final states by Dill and Dehmer<sup>124</sup> and Davenport<sup>125</sup>. This method adopts two approximations in calculations dealing with relatively large clusters, namely, one to the exchange potential in the HF Hamiltonian, and the other one to the overall cluster potential; the former is approximated by an average potential, determined by the total charge density, while the latter is approximated by a “muffin-tin” form. Other methods rely on time-dependent density functional theory (TDDFT); among the several examples, the complex

propagator approach (CPP) by Norman and co-workers has been particularly well exploited<sup>126,127</sup>. Within this approach, the scanning of the imaginary part of the complex electric dipole polarizability over the investigated frequency region gives the absorption cross section of the core excitation process. Another well-suited method is TDDFT with a restricted excitation manifold<sup>128,129</sup> in a similar spirit to that of the core–valence separation (CVS) approximation<sup>111</sup>. The latter is used in most wave function-based methods for calculating X-ray excitations, and has been first implemented within the second-order Algebraic Diagrammatic Construction (ADC(2)) method<sup>130</sup>. Wenzel and co-workers have recently applied the CVS approximation for the ADC(2)-x and ADC(3) hierarchy of methods<sup>131</sup> based on the intermediate state representation variant of the ADC formalism<sup>132,133</sup>. The calculation of core-excitation spectra has been also performed within coupled cluster (CC) theory; indeed, it is well known that CC methods are able to provide particularly accurate results for core excitation energies and intensities, especially for small molecules<sup>134-136</sup>. In these methods, the computation of excited states requires the extraction of the eigenvalues of the Jacobian; the full spectrum can be obtained through the diagonalization of the Jacobian matrix which can be constructed by means of an asymmetric Lanczos algorithm.<sup>134,135</sup>

## 2.9 Computational details

The commercial molecular DFT code ADF (*Amsterdam Density Functional*)<sup>137,138</sup> has been employed for the computation of NEXAFS spectra. To expand the KS orbitals, the ADF program uses Slater Type Orbitals (STO) atomic basis sets. The radial part of the STO basis functions consists of an exponential part and a polynomial pre-factor:

$$\Psi_{n,l,\alpha}^{STO}(r, \theta, \varphi) = Nr^{n-l-1}e^{-\alpha r} [r^l Y_{l,m}(\theta, \varphi)] \quad (2.164)$$

where N is a normalization constant,  $n$ ,  $l$  and  $m$  are the common atomic quantum numbers, while the exponent  $\alpha$  is variationally optimized for each atom.

KS MOs are built through a linear combination of basis functions; KS equations (Eq. 2.66) are hence rewritten as an eigenvalue algebraic problem. Preliminary convergence tests can be usually performed to choose the best basis set to properly reproduce the experimental measurements.

To fit the electronic density, a second auxiliary basis set is employed in ADF. Indeed, to minimize the computational effort of the Coulomb term in the Hamiltonian, one can approximate the true density (*i.e.*, a sum of products of basis functions) with a linear combination of the fit functions  $f_i(\mathbf{r})$ , such that:

$$\tilde{\rho}(\mathbf{r}) = \sum_i a_i f_i(\mathbf{r}) \quad (2.165)$$

where  $a_i$  are the fit coefficients.

The fit functions  $f_i(\mathbf{r})$  can be extracted from a data base included in ADF, whereas the coefficients  $a_i$  are obtained through a least-squares minimization of the error:

$$\int [\rho(\mathbf{r}) - \tilde{\rho}(\mathbf{r})]^2 d\mathbf{r} \quad (2.166)$$

with the constraint  $\int \tilde{\rho}(\mathbf{r}) d\mathbf{r} = N$ .

Once the electronic density is fitted with the outlined procedure, the Coulomb term of the KS Hamiltonian matrix is computed through Gaussian quadrature between the basis function pair and the Coulomb potential deriving from  $\tilde{\rho}(\mathbf{r})$ .

The fit set and the accuracy of the fit approximation both play a fundamental role, similar to that of the basis set: it is well known that too few functions (or badly chosen function characteristics) yield less accurate results. However, unlike the basis set, the size of the fit set does not drastically affect the computational effort.

In the computation of the NEXAFS spectra, a rich basis set including polarization and diffuse functions is usually used only for the excited atom. In this thesis work, the even tempered ET-QZ4P-2DIFFUSE basis set of the ADF database has been employed for the excited atoms of all investigated systems. It consists of four atomic functions for each valence atomic orbital and three polarization functions with the addition of two diffuse functions. Polarization functions are characterized by an angular momentum quantum number higher than one unit compared to the one of the outer valence sub-shell, and are fundamental in the description of the charge density's polarization upon the formation of highly directional chemical bonds in molecules. Diffuse functions have principal quantum number higher than one unit (or more) compared to the atomic valence orbitals, and are necessary to describe the high-energy transitions toward Rydberg states. The non-excited atoms have been treated with a lower accuracy through a triple  $\zeta$  polarized (TZP) basis set. The latter consists of two atomic functions for each core orbital, and three atomic

functions for each valence orbital, with the addition of a polarization function; the TZP basis set has been also adopted in the geometry optimization computations. Furthermore, the Frozen Core (FC) technique has been employed for the non-excited atoms in NEXAFS calculations to ensure the localization of the core-hole: it consists in treating explicitly only outer level electrons, and keeping frozen the innermost (core) atomic shells. The basis set employed in NEXAFS calculations of each system investigated in this thesis work will be described in the Computational Details section of the selected research outputs.

## Bibliography

---

- <sup>1</sup> J. Stöhr, *NEXAFS Spectroscopy*; Springer Verlag: Berlin, Heidelberg, Germany, and New York, **1992**
- <sup>2</sup> L.G.M. Pettersson, H. Ågren, Y. Luo, L. Triguero, *Surf. Sci.* **1998**, *408*, 1-20
- <sup>3</sup> M. Stener, A. Lisini, P. Decleva, *J. Electron Spectrosc. Relat. Phenom.* **1994**, *69*, 197-206
- <sup>4</sup> R.J. Sundberg, *The Chemistry of Indoles*; Academic Press: New York, **1996**
- <sup>5</sup> D. Deletioğlu, E. Hasdemir, A.O. Solak, Z. Üstündağ, R. Güzel, *Thin Solid Films*, **2010**, *519*, 784–789
- <sup>6</sup> M. Ghita, D.W. Arrigan, *Electroanalysis*, **2004**, *16*, 979–987
- <sup>7</sup> C. Duan, K. Gao, F.J.M. Colberts, F. Liu, S.C.J. Meskers, M.M. Wienk, R.A.J. Janssen, *Adv. Energy Mater.*, **2017**, *7*, 1700519, 1-11
- <sup>8</sup> S.H. Jeong, J.Y. Lee, *J. Mater Chem.*, **2011**, *21*, 14604-14609
- <sup>9</sup> D. Kim, S. Salman, V. Coropceanu, E. Salomon, A.B. Padmaperuma, L.S. Sapochak, A. Kahn, J.-L. Bredas, *J. Chem. Mater.*, **2010**, *22*, 247-254
- <sup>10</sup> Y.-S. Tsai, L.-A. Hong, F.-S. Juang, C.-Y. Chen, *J. Lumin.*, **2014**, *153*, 312–316
- <sup>11</sup> M. Chapran, E. Angioni, N.J. Findlay, B. Breig, V. Cherpak, P. Stakhira, T. Tuttle, D. Volyniuk, J.V. Grazulevicius, Y.A. Nastishin, et al. *ACS Appl. Mater. Interfaces*, **2017**, *9*, 4750–4757
- <sup>12</sup> A.B. Padmaperuma, L.S. Sapochak, P.E. Burrows, *Chem. Mater.* **2006**, *18*, 2389–2396
- <sup>13</sup> D.G. Gilheany, *Chem. Rev.* **1994**, *94*, 1339–1374
- <sup>14</sup> Z. Zhang et al., *Angew. Chem.*, Int. Ed. **2016**, *55*, 5206–5209
- <sup>15</sup> N. Coppède et al., *Org. Electron.*, **2016**, *32*, 15–20
- <sup>16</sup> Y.-S. Huang et al., *Appl. Phys. Lett.*, **2002**, *80*, 2782-2784
- <sup>17</sup> M.G. Walter, A.B. Rudine, C.C. Wamser, *J. Porphyrins Phthalocyanines*, **2010**, *14*, 759-792
- <sup>18</sup> Y. Wei, S.W. Robey, J.E. Reutt-Robey, *J. Phys. Chem. C* **2008**, *112*, 18537–18542
- <sup>19</sup> X. Liu, Y. Wei, J.E. Reutt-Robey, S.W. Robey, *J. Phys. Chem. C* **2014**, *118*, 3523–3532
- <sup>20</sup> L. Fernández, S. Thussing, A. Mänz, G. Witte, A.X. Brion-Rios, P. Cabrera-Sanfelix, D. Sanchez-Portal, P. Jakob, *J. Phys. Chem. C* **2017**, *121*, 1608–1617
- <sup>21</sup> R. Nishiyabu, Y. Kubo, T.D. James, J.S. Fossey, *Chem. Commun.*, **2011**, *47*, 1124-1150
- <sup>22</sup> S. Clair, M. Abel, L. Porte, *Chem. Commun.*, **2014**, *50*, 9627-9635
- <sup>23</sup> M. Stredansky, A. Sala, T. Fontanot, R. Costantini, C. Africh, G. Comelli, L. Floreano, A. Morgante, A. Cossaro, *Chem. Commun.*, **2018**, *54*, 3971-3973

- 
- <sup>24</sup> D. Toffoli M. Stredansky, Z. Feng, G. Balducci, S. Furlan, M. Stener, H. Ustunel, D. Cvetko, G. Kladnik, A. Morgante, A. Verdini, C. Dri, G. Comelli, G. Fronzoni, A. Cossaro, *Chem. Sci.* **2017**, *8*, 3789-3798
- <sup>25</sup> J.C. Slater, *Adv. Quantum Chem.* **1972**, *6*, 1–92
- <sup>26</sup> M.E. Casida, *Time-Dependent Density Functional Response Theory for Molecules. Recent Advances in Density Functional Methods* **1995**, *1*, 155-192
- <sup>27</sup> B.O. Roos, *Adv. Chem. Phys.* **1987**, *69*, 399-445; H.-J. Werner, P.J. Knowles, *J. Chem. Phys.* **1985**, *82*, 5053-5063
- <sup>28</sup> P.J. Knowles, H.-J. Werner, *Chem. Phys. Lett.* **1985**, *115*, 259-267
- <sup>29</sup> G. Fronzoni, G. Balducci, R. De Francesco, M. Romeo, M. Stener, *J. Phys. Chem. C* **2012**, *116*, 18910–18919
- <sup>30</sup> R. De Francesco, M. Stener, G. Fronzoni, *J. Phys. Chem.* **2009**, *11*, 1146–1151
- <sup>31</sup> G. Fronzoni, R. De Francesco, M. Stener, M. Causà, *J. Phys. Chem. B* **2006**, *110*, 9899-9907
- <sup>32</sup> R. De Francesco, M. Stener, M. Causà, D. Toffoli, G. Fronzoni, *Phys. Chem. Chem. Phys.* **2006**, *8*, 4300-4310
- <sup>33</sup> R.H. Messner, *Z. Phys.* **1933**, *85*, 727-740
- <sup>34</sup> T. Magnusson, *Nova Acta Regiae Soc. Sci. Upsal.* **1929**, *11*, n°3
- <sup>35</sup> A. Bianconi, in *X-ray absorption: principles, applications, techniques of EXAFS, SEXAFS and XANES*, Koningsberger D. C., Prins R. (Eds.), John Wiley, New York, **1988**, p. 573
- <sup>36</sup> R.A. Buckingham, *In Quantum Theory I. Elements*, ed. by D.R. Bates (Academic, New York **1961**), p. 147
- <sup>37</sup> A. Bianconi, *Appl. Surf. Sci.* **1980**, *6*, 392-418
- <sup>38</sup> W. Kossel, *Z. Phys.* **1920**, *1*, 119-134; *ibid.* **1920**, *2*, 470-478
- <sup>39</sup> J. Bergengren, *Z. Phys.* **1920**, *3*, 247-249
- <sup>40</sup> A.E. Lindh, *Z. Phys.* **1921**, *6*, 303-310
- <sup>41</sup> R. de L. Kronig, *Z. Phys.*, **1931**, *70*, 317-323; *ibid.* **1932**, *75*, 468-475
- <sup>42</sup> J. Stohr, R. Jaeger, *Phys. Rev. B* **1982**, *26*, 4111-4131
- <sup>43</sup> J.L. Dehmer, A.C. Parr, S.H. Southworth, *In Handbook on Synchrotron Radiation*, Vol. 2, ed. by G.V. Marr (North-Holland, Amsterdam **1987**), p. 241
- <sup>44</sup> G. Hähner, *Chem. Soc. Rev.* **2006**, *35*, 1244-1255
- <sup>45</sup> F. Sette, J. Stohr, A.P. Hitchcock, *J. Chem. Phys.* **1984**, *81*, 4906-4914
- <sup>46</sup> J. Stohr, K. Baberschke, R. Jaeger, T. Treichler, S. Brennan, *Phys. Rev. Lett.* **1981**, *47*, 381-384
- <sup>47</sup> G. Blyholder, *J. Phys. Chem.* **1964**, *68*, 2772-2777

- 
- <sup>48</sup> G. Blyholder, M.E. Allen, *J. Am. Chem. Soc.* **1969**, *91*, 3158-3162
- <sup>49</sup> E.W. Plummer, W.R. Salaneck, J.S. Miller, *Phys. Rev. B* **1978**, *18*, 1673-1701
- <sup>50</sup> H.-J. Freund, E.W. Plummer, *Phys. Rev. B*, 1981, *23*, 4859-4878
- <sup>51</sup> E.W. Plummer, W.K. Ford, W. Eberhard, R. Messmer, H.-J. Freund, *Surf. Sci.* **1985**, *158*, 58-83
- <sup>52</sup> A. Yamamoto, *Organotransition Metal Chemistry: Fundamental Concepts and Applications* (Wiley, New York **1986**)
- <sup>53</sup> M.I.S. Dewar, *Bull. Soc. Chem.* 1951, Fr. C71
- <sup>54</sup> J. Chatt, L.A. Duncanson, *J. Chem. Soc.* **1953**, 2939-2947
- <sup>55</sup> K. Siegbahn, *ESCA: Atomic, Molecular and Solid State Structure Studied by Means of Electron Spectroscopy* **1967**, vol. 20
- <sup>56</sup> "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt" *Ann. Phys.* (Berlin) **1905**, *17*, 132-148
- <sup>57</sup> T. Koopmans, *Physica* **1934**, *1*, 104-113
- <sup>58</sup> J.F. Mulder, C.D. Wagner, W.M. Riggs, L.E. Davis, *Handbook of X-ray Photoelectron Spectroscopy*, **1965**
- <sup>59</sup> S. Hüfner, *Photoelectron Spectroscopy*, Springer-Verlag Berlin Heidelberg GmbH, ISBN 9783642075209
- <sup>60</sup> P.J. Stephens, F.J. Devlin, C.F. Chabalowski, M.J. Frisch, *J. Phys. Chem.*, **1994**, *98(45)*, 11623-11627
- <sup>61</sup> A. Szabo, N.S. Ostlund, *Modern Quantum Chemistry*, McGraw-Hill, **1982**
- <sup>62</sup> P. Politzer, A.-A. Fakhri, *Th. Chem. Acc.* **1998**, *99* (2), 83-87
- <sup>63</sup> S. Hamel, P. Duffy, M.E. Casida, D.R. Salahub, *J. Electron Spectrosc. Relat. Phenom.* 2002, *123* (2-3), 345-363
- <sup>64</sup> W.J. Hehre, L. Radom, P.v.R. Schleyer, J.A. Pople, *Ab initio molecular orbital theory*, p. 24, New York, **1986**
- <sup>65</sup> C.D. Sherrill, H.F. Schaefer, *Advances in Quantum Chemistry*, volume 34, pp. 143-269, Academic, New York, **1999**
- <sup>66</sup> M. Born, J.R. Oppenheimer, *Ann. Physik* **1927**, *389*, 457-484
- <sup>67</sup> M.W. Schmidt, M.S. Gordon, *Ann. Rev. Phys. Chem.* **1998**, *49*, 233-266
- <sup>68</sup> B.O. Roos, P.R. Taylor, P.E.M. Siegbahn, *Chem. Phys.* **1980**, *48*, 157-173
- <sup>69</sup> R.G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, **1989**
- <sup>70</sup> L.H. Thomas, *Proc. Cambridge Phil. Soc.* **1927**, *23*, 542-548



- 
- <sup>71</sup> E. Fermi, *Rend. Accad. Naz. Lincei* **1927**, 6, 602–607
- <sup>72</sup> E.H. Lieb, B. Simon, *Phys. Rev. Lett.* **1973**, 31, 681–683
- <sup>73</sup> E. Teller, *Rev. Mod. Phys.* **1962**, 34, 627–631
- <sup>74</sup> P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, 136, B:864-B871
- <sup>75</sup> M. Levi, J. Perdew, *The Constrained Search Formulation of Density Functional Theory*, Springer US, Boston, MA, **1985**, pp. 11-30
- <sup>76</sup> C. Ullrich, *Time-Dependent Density Functional Theory: Concepts and Applications*, Oxford Graduate Texts, OUP Oxford, **2012**
- <sup>77</sup> W. Kohn, L.J. Sham, *Phys. Rev.* **1965**, 140, A1133-A1138
- <sup>78</sup> S.H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, 58, 1200–1211
- <sup>79</sup> D.M. Ceperley, J.B. Alder, *Phys. Rev. Lett.* **1980**, 45, 566–569
- <sup>80</sup> J.P. Perdew, A. Zunger, *Phys. Rev. B: Condens. Matter* **1981**, 23, 5048–5079
- <sup>81</sup> J.P. Perdew, Y. Wang, *Phys. Rev. B: Condens. Matter* **1992**, 45, 13244–13249
- <sup>82</sup> R. van Leeuwen, E.J. Baerends, *Phys. Rev. A: At., Mol., Opt. Phys.* **1994**, 49, 2421–2431
- <sup>83</sup> A.D. Becke, *Phys. Rev. A* **1988**, 38, 3098-3100
- <sup>84</sup> C. Lee, W. Yang, R.G. Parr, *Phys. Rev. B* **1988**, 37, 785-789
- <sup>85</sup> J.P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, 77, 3865-3868
- <sup>86</sup> J.P. Perdew, *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, 33, 8822–8824
- <sup>87</sup> Y. Zhao, D.G. Truhlar, *J. Chem. Phys.* **2006**, 125, 194101-1; 194101-18
- <sup>88</sup> J.P. Perdew, S. Kurth, A. Zupan, P. Blaha, *Phys. Rev. Lett.* **1999**, 82, 2544-2547
- <sup>89</sup> J. Tao, J.P. Perdew, V.N. Staroverov, G.E. Scuseria, *Phys. Rev. Lett.* **2003**, 91, 146401–146404
- <sup>90</sup> J.P. Perdew, V.N. Staroverov, J. Tao, G.E. Scuseria, *Phys. Rev. A* **2008**, 78, 052513–052525
- <sup>91</sup> L.A. Constantin, J.M. Pitarke, J.F. Dobson, A. Garcia-Lekue, J.P. Perdew, *Phys. Rev. Lett.* **2008**, 100, 036401–036404
- <sup>92</sup> J.P. Perdew, A. Ruzsinszky, L.A. Constantin, J. Sun, G.I. Csonka, *J. Chem. Theory Comput.* **2009**, 5, 902–908
- <sup>93</sup> A.D. Becke, *J. Chem. Phys.* **1993**, 98(2), 1372-1377
- <sup>94</sup> C. Lee, W. Yang, R.G. Parr, *Phys. Rev. B*, **1988**, 37, 785-789
- <sup>95</sup> O. Goscinski, B.T. Pickup, G. Purvis, *Chem. Phys. Lett.*, **1973**, 22, 167-171
- <sup>96</sup> G. Howat, O. Goscinski, *Chem. Phys. Lett.*, **1975**, 30, 87-90
- <sup>97</sup> A.R. Williams, R.A. de Groot, C.B. Sommers, *J. Chem. Phys.*, **1975**, 63, 628-631
- <sup>98</sup> L. Triguero, L.G.M. Pettersson, H. Ågren, *Phys. Rev. B*, **1998**, 58, 8097-8110
- <sup>99</sup> J.C. Slater, K.H. Johnson, *Phys. Rev. B*, **1972**, 5, 844-853

- 
- <sup>100</sup> R. Hoffmann, *A chemist's view of bonding in extended structures* (VCH Publishers, New York, **1988**)
- <sup>101</sup> R. Bauernschmitt, R. Ahlrichs, *Chem. Phys. Lett.* **1996**, *256*, 454–464
- <sup>102</sup> F. Wang, T. Ziegler, *Mol. Phys.* **2004**, *102*, 2585–2595
- <sup>103</sup> S.J.A. van Gisbergen, J.G. Snijders, E.J. Baerends, *Comput. Phys. Commun.* **1999**, *118*, 119–138
- <sup>104</sup> G. Fronzoni, M. Stener, P. Decleva, F. Wang, T. Ziegler, E. van Lenthe, E.J. Baerends, *Chem. Phys. Lett.* **2005**, *416*, 56–63
- <sup>105</sup> E. van Lenthe, E.J. Baerends, J.G. Snijders, *J. Chem. Phys.* **1993**, *99*, 4597–4610
- <sup>106</sup> E. van Lenthe, E.J. Baerends, J.G. Snijders, *J. Chem. Phys.* **1994**, *101*, 9783–9792
- <sup>107</sup> E. van Lenthe, A. Ehlers, E.J. Baerends, *J. Chem. Phys.* **1999**, *110*, 8943–8953
- <sup>108</sup> G. Diercksen, R. McWeeny, *J. Chem. Phys.* **1966**, *44*, 3554–3560
- <sup>109</sup> E.K.U. Gross, W. Kohn, *Phys. Rev. Lett.* **1985**, *55*, 2850–2852
- <sup>110</sup> E.R. Davidson, *J. Comput. Phys.* **1975**, *17*, 87–94
- <sup>111</sup> L.S. Cederbaum, W. Domcke, J. Schirmer, *Phys. Rev. A: At., Mol., Opt. Phys.* **1980**, *22*, 206–222
- <sup>112</sup> A. Barth, L.S. Cederbaum, *Phys. Rev. A: At., Mol., Opt. Phys.* **1981**, *23*, 1038–1061
- <sup>113</sup> P. Decleva, G. Fronzoni, A. Lisini, M. Stener, *Chem. Phys.* **1994**, *186*, 1–16
- <sup>114</sup> J.B. Foresman, M. Head-Gordon, J.A. Pople, M.J. Frisch, *J. Phys. Chem.* **1992**, *96*, 135–149
- <sup>115</sup> G. Fronzoni, O. Baseggio, M. Stener, W. Hua, G. Tian, Y. Luo, B. Apicella, M. Alfé, M. de Simone, A. Kivimäki, M. Coreno, *J. Chem. Phys.* **2014**, *141*, 044313-1; 044313-9
- <sup>116</sup> A. Baiardi, M. Mendolicchio, V. Barone, G. Fronzoni, G.A. Cardenas Jimenez, M. Stener, C. Grazioli, M. de Simone, M. Coreno, *J. Chem. Phys.* **2015**, *143*, 204102-1; 204102-12
- <sup>117</sup> C. Grazioli, O. Baseggio, M. Stener, G. Fronzoni, M. de Simone, M. Coreno, A. Guarnaccio, A. Santagata, M. D'Auria, *J. Chem. Phys.* **2017**, *146*, 054303-1; 054303-9
- <sup>118</sup> C. Ehlert, T. Klamroth, *J. Comput. Chem.* **2017**, *38*, 116–126
- <sup>119</sup> O. Baseggio, D. Toffoli, M. Stener, G. Fronzoni, M. de Simone, C. Grazioli, M. Coreno, A. Guarnaccio, A. Santagata, M. D'Auria, *J. Chem. Phys.* **2017**, *147*, 244301-1; 244301-13
- <sup>120</sup> G. Fronzoni, R. De Francesco, M. Stener, P. Decleva, *J. Chem. Phys.* **2007**, *126*, 134308-1; 134308-10
- <sup>121</sup> F. Wang, T. Ziegler, E. van Lenthe, S. van Gisbergen, E.J. Baerends, *J. Chem. Phys.* **2005**, *122*, 204103-1; 204103-13
- <sup>122</sup> P. Schwerdtfeger, *Chem. Phys. Chem* **2011**, *12*, 3143–3155
- <sup>123</sup> K.H. Johnson, *Advances in Quantum Chemistry*, Vol. 7, ed. by P.-O. Loewdin (Academic, New York **1973**) p. 143

- 
- <sup>124</sup> D. Dill, J.L. Dehmer, *J. Chem. Phys.* **1974**, *61*, 692-699
- <sup>125</sup> J.W. Davenport, *Phys. Rev. Lett.* **1976**, *36*, 945-949
- <sup>126</sup> U. Ekström, P. Norman, V. Carravetta, H. Ågren, *Phys. Rev. Lett.* **2006**, *97*, 143001-1; 143001-4
- <sup>127</sup> U. Ekström, P. Norman, *Phys. Rev. A: At., Mol., Opt. Phys.* **2006**, *74*, 042722-1; 042722-7
- <sup>128</sup> N.A. Besley, F.A. Asmuruf, *Phys. Chem. Chem. Phys.* **2010**, *12*, 12024–12039
- <sup>129</sup> M. Stener, G. Fronzoni, M. de Simone, *Chem. Phys. Lett.* **2003**, *373*, 115–123
- <sup>130</sup> J. Schirmer, *Phys. Rev. A: At., Mol., Opt. Phys.* **1982**, *26*, 2395–2416
- <sup>131</sup> A.B. Trofimov, G. Stelter, J. Schirmer, *J. Chem. Phys.* **1999**, *111*, 9982-9999
- <sup>132</sup> M. Wormit, D.R. Rehn, P.H. Harbach, J. Wenzel, C.M. Krauter, E. Epifanovsky, A. Dreuw, *Mol. Phys.* **2014**, *112*, 774–784
- <sup>133</sup> J. Wenzel, M. Wormit, A. Dreuw, *J. Comput. Chem.* **2014**, *35*, 1900–1915
- <sup>134</sup> S. Coriani, T. Fransson, O. Christiansen, P. Norman, *J. Chem. Theory Comput.* **2012**, *8*, 1616–1628
- <sup>135</sup> S. Coriani, O. Christiansen, T. Fransson, P. Norman, *Phys. Rev. A: At., Mol., Opt. Phys.* **2012**, *85*, 022507-1; 022507-8
- <sup>136</sup> T. Fransson, S. Coriani, O. Christiansen, P. Norman, *J. Chem. Phys.* **2013**, *138*, 124311-1; 124311-12
- <sup>137</sup> C. Fonseca Guerra, J.G. Snijders, G. te Velde, E.J. Baerends, *Theor. Chem. Acc.* **1998**, *99*, 391–403
- <sup>138</sup> E.J. Baerends, D.E. Ellis, P. Ros, *Chem. Phys.* **1973**, *2*, 41–51

The following discussion reports selected results obtained during the three-year period of PhD which have been briefly described in the Overview. Conclusions and perspectives on the research activity will be finally provided.

### List of publications

- 1) O. Plekan, H. Sa'adeh, A. Ciavardini, C. Callegari, G. Cautero, C. Dri, M. Di Fraia, K.C. Prince, R. Richter, R. Sergo, L. Stebel, M. Devetta, D. Faccialà, C. Vozzi, L. Avaldi, P. Bolognesi, M.C. Castrovilli, D. Catone, M. Coreno, F. Zuccaro, **E. Bernes**, G. Fronzoni, D. Toffoli, A. Ponzi, *J. Phys. Chem. A* **2020**, *124*, 4115-4127;
- 2) D. Toffoli, A. Guarnaccio, C. Grazioli, T. Zhang, F. Johansson, M. de Simone, M. Coreno, A. Santagata, M. D'Auria, C. Puglia, **E. Bernes**, M. Stener, G. Fronzoni, *J. Phys. Chem. A* **2018**, *122*, 44, 8745-8761;
- 3) A. Guarnaccio, T. Zhang, C. Grazioli, F. Johansson, M. Coreno, M. de Simone, G. Fronzoni, D. Toffoli, **E. Bernes**, C. Puglia, *J. Phys. Chem. C* **2020**, *124*, 9774-9786;
- 4) **E. Bernes**, G. Fronzoni, M. Stener, A. Guarnaccio, T. Zhang, C. Grazioli, F. Johansson, M. Coreno, M. de Simone, C. Puglia, D. Toffoli, *J. Phys. Chem. C* **2020**, *124*, 14510–14520;
- 5) D. Toffoli, A. Ponzi, **E. Bernes**, M. de Simone, C. Grazioli, M. Coreno, M. Stredansky, A. Cossaro, G. Fronzoni, *J. Chem. Phys.* **2019**, *151*, 134306-1; 134306-8.

# Experimental and Theoretical Photoemission Study of Indole and Its Derivatives in the Gas Phase

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

Cite This: *J. Phys. Chem. A* 2020, 124, 4115–4127

Read Online

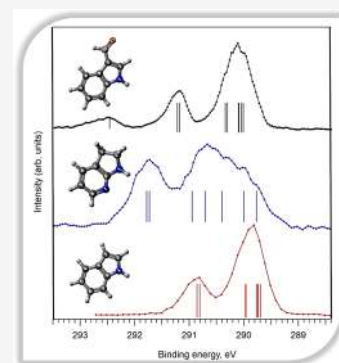
ACCESS |

Metrics & More

Article Recommendations

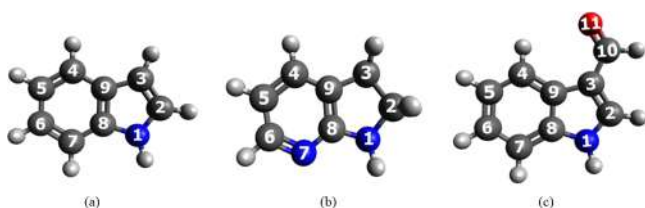
Supporting Information

**ABSTRACT:** The valence and core-level photoelectron spectra of gaseous indole, 2,3-dihydro-7-azaindole, and 3-formylindole have been investigated using VUV and soft X-ray radiation supported by both an ab initio electron propagator and density functional theory calculations. Three methods were used to calculate the outer valence band photoemission spectra: outer valence Green function, partial third order, and renormalized partial third order. While all gave an acceptable description of the valence spectra, the last method yielded very accurate agreement, especially for indole and 3-formylindole. The carbon, nitrogen, and oxygen 1s core-level spectra of these heterocycles were measured and assigned. The double ionization appearance potential for indole has been determined to be  $21.8 \pm 0.2$  eV by C 1s and N 1s Auger photoelectron spectroscopy. Theoretical analysis identifies the doubly ionized states as a band consisting of two overlapping singlet states and one triplet state with dominant configurations corresponding to holes in the two uppermost molecular orbitals. One of the singlet states and the triplet state can be described as consisting largely of a single configuration, but other doubly ionized states are heavily mixed by configuration interactions. This work provides full assignment of the relative binding energies of the core level features and an analysis of the electronic structure of substituted indoles in comparison with the parent indole.



## INTRODUCTION

Indole (I), 2,3-dihydro-7-azaindole (7-AI), and 3-formylindole (3-FI) (see Figure 1) heterocycles are prevalent substructures



**Figure 1.** Schematic chemical structures of the investigated indoles: (a) indole (I), (b) 2,3-dihydro-7-azaindole (7-AI), and (c) 3-formylindole with the trans orientation of the formyl group (3-FI). The C (black), N (blue), and O (red) atoms are labeled.

in naturally occurring and synthetic molecules playing key roles in chemistry and biology.<sup>1</sup> Indole derivatives possess unique biological characteristics such as antioxidant, anticancer, antibacterial, antifungal, anti-inflammatory, antiviral, anticonvulsant, and antihypertensive properties among other applications. In addition, indole-based chromophores form the basic

building block units of the eumelanin pigments that occur frequently in nature and serve to protect organisms from potentially damaging effects of UV light.<sup>2</sup> 7-AI, whose molecular structure is shown in Figure 1b, has received considerable attention since the corresponding dimer has been recognized as a simple model for the hydrogen-bonded base pairs of DNA and could provide information on the possible role of tautomerism in mutation.<sup>3–5</sup> 3-Formylindole is a heteroarenealdehyde, that is, the hydrogen at position 3 of indole has been replaced by a formyl group (see Figure 1c), and is known as a biologically active metabolite.<sup>6</sup>

Additionally, in recent years, indoles have received increasing attention as coating materials, because they are easily electrografted to surfaces, forming conductive films.<sup>7</sup> Indole-coated surfaces have widespread applications and can

Received: March 27, 2020

Revised: April 24, 2020

Published: April 24, 2020



be used as selective electrodes sensitive to various cationic and anionic inorganic species, as a biosensor for biological molecules, or as a protection of metallic surfaces against corrosion.<sup>8,9</sup> According to the literature, many intrinsic properties (for example, the polymerization mechanism) of organic heterocycle compounds are masked by their environment or by their interactions with it.<sup>10,11</sup> Thus, gas-phase data provide a better understanding of the properties of indole and its derivatives in the absence of perturbations due to their interactions with different surfaces.

Because of their rich chemistry and biological activity, indoles have been extensively studied using various spectroscopies. Experimental studies of electronic properties of indole and its derivatives have employed optical (including vibrationally and rotationally resolved) methods<sup>12–19</sup> as well as time-resolved ion and photoelectron spectroscopy<sup>20–24</sup> as these powerful techniques probe the properties of greatest interest. Theoretical studies of these heterocycles have been mainly focused on optimization of the structure and calculations of vibrational spectra using the *ab initio* Hartree–Fock (HF), density functional theory (DFT), and MP2 methods.<sup>25–29</sup>

Recently, the photoinduced structural transformations of indole, 7-azaindole, and 3-formylindole isolated in cryogenic noble-gas matrices were studied by Nowak and co-workers.<sup>30,31</sup> It was found that the tendency of indole and other indole derivatives to undergo hydrogen-atom transfer from N1 to C3 seems to be strongly related with the formation of C3-centered radicals, which facilitate the reattachment of the labile hydrogen atom at this position. The authors<sup>31</sup> have observed that the pattern of photochemical transformations found for 7-AI isolated in solid *n*-H<sub>2</sub> and irradiated with UV ( $\lambda > 270$  nm) light, is significantly different from that observed for the compound trapped in solid Ar. While the C3H tautomer and 7-azaindolyl radical were photogenerated in both environments (solid Ar and solid *n*-H<sub>2</sub>), the N7H tautomer was not photoproducted from 7-AI isolated in solid *n*-H<sub>2</sub>. Additionally, the conformational 1H-*trans*  $\leftrightarrow$  1H-*cis* phototransformation by rotation of the aldehyde group in 3-formylindole has been studied as well.<sup>30</sup> It was proposed that the 1H-*trans* isomer of 3-FI is the most populated form of the compound isolated in Ar matrices, whereas the 1H-*cis* isomer could also be populated but with a considerably lower population.

The photoelectron spectrum (PES) of indole was measured many years ago by Eland.<sup>32</sup> Later the PES of indole in the outer valence region was studied with a He I (21.2 eV) source in order to determine whether any correlation between the ionization potential and drug activity exists.<sup>33</sup> Kovač et al. recorded He I PES of 10 different nitrogen-containing heterocycles, and assignment of the results was provided on the basis of molecular orbital calculations.<sup>34</sup> In the late 1980s, multiphoton ionization photoelectron spectroscopy was used in order to obtain information about the adiabatic and vertical ionization potentials of 7-AI in a supersonic jet.<sup>35</sup> Recently, a comprehensive electronic-structure analysis of parental structures of BN indoles (bicyclic aromatic heterocycles in which a 1,3,2-diazaboroline is fused to a benzene ring) in direct comparison to the parent indole was carried out by Chrostowska et al.<sup>36</sup> using a combined UV-PES and computational chemistry approach.

An inner-shell photoionization study of a gas-phase indole was published by Kierspel and co-authors.<sup>37</sup> Detailed photoionization and photofragmentation spectra of indole upon single-photon inner-shell ionization at a photon energy of 420

eV were recorded. This photon energy was chosen so that indole could be locally ionized at its nitrogen or carbon atoms. In addition, electrons and ions were measured in coincidence in a velocity-map-imaging mode to extract 2D and 3D velocity vectors of the charged particles. It was found that the fragmentation channels depend on the ionized electronic states, i.e., the potential energy surfaces, whereas the observed velocities of the fragments are not strongly dependent on these chemical details.<sup>37</sup> However, it appears that up to now experimental evidence concerning the full valence and core electronic structure for derivatives of indole under isolated conditions is still limited. To the best of our knowledge, complete electronic structure investigations obtained by soft X-ray photoemission and near edge X-ray absorption fine-structure (NEXAFS) spectroscopy only exist for gaseous 3-methylindole.<sup>38,39</sup> The valence molecular orbitals and core levels of tryptamine and tryptophol (both contain indole rings but differ in having amino versus hydroxyl terminations of the side chain) in the gas phase have been studied using X-ray photoelectron spectroscopy (XPS) and theoretical methods.<sup>40</sup>

Considering the general importance of these nitrogen-containing heterocycles, an accurate knowledge of their electronic structure becomes crucial. In the present work, we provide a comprehensive electronic structure analysis for two indole derivatives - in direct comparison to the parent indole - by combining valence band (VB) and XPS with theoretical calculations. VB spectra allow us to experimentally determine the ionization energies of molecules that can be correlated to the energies of occupied molecular orbitals, while XPS provides localized ionization of a specific atom, and chemical information about that atom.

## ■ EXPERIMENTAL AND THEORETICAL METHODS

**Theoretical Methods.** Geometry optimizations for the gas-phase molecules were carried out at the DFT B3LYP/aug-cc-pVTZ level by using the Gaussian09 program,<sup>41</sup> and the optimized structures presented in Figure 1 were used for all subsequent calculations. The complete charge density maps of the outermost molecular orbitals (MOs) of I, 7-AI, and 3-FI within the Hartree–Fock/cc-pVTZ model are reported in Figures S1–S3 of the Supporting Information.<sup>42</sup> Note that, in the case of the 3-formylindole molecule, all calculations were performed for its more stable 1H-*trans* (0.0 kJ/mol) conformer rather than the less stable 1H-*cis* (5.1 kJ/mol) conformer.<sup>30</sup>

To a first approximation, the VB ionization spectrum can be described by exploiting Koopmans' theorem (KT). This establishes a direct correlation between the single bands of the spectrum and the MO energies. KT values corresponding to the vertical ionization potential (IP) are reported in Tables S1–S3 of the Supporting Information.<sup>42</sup>

To obtain a more accurate description of the valence spectra, vertical IPs have been calculated with three different *ab initio* electron propagator (EP) methods. The IPs are obtained as poles of the EP, which in turn correspond to the eigenvalues of a Dyson equation, which must be solved self-consistently.<sup>43</sup> These methods are only applicable in instances where KT provides a reasonable description of initial and final states (quasiparticle approximation or an orbital picture of ionization), and the Dyson orbitals are therefore proportional to the HF orbitals. Diagonal approximations usually fail in the presence of strong relaxation effects and cannot describe the strong redistribution of intensity from the main lines to satellite states, which is usually observed in inner-valence

**Table 1. Experimental and Calculated Outer Valence Band Binding Energies of Indole, 2,3-Dihydro-7-azaindole, and 3-Formylindole**

compounds	experimental BEs (eV), present work, $\pm 0.1$ eV	calculated BEs (eV) using the P3+/cc-pVTZ approach / molecular orbital (type)	published BEs (eV) <sup>33,35,36</sup>
indole	7.72 <sup>a</sup>		
	7.77 (shoulder)	7.76 <sup>a</sup> (shoulder)	7.75 <sup>33</sup>
	7.90		7.92; <sup>33</sup> 7.9 <sup>36</sup>
	8.07 (shoulder)	8.05 <sup>a</sup>	8.08 <sup>33</sup>
	8.32	8.35 <sup>a</sup>	8.37; <sup>33</sup> 8.5 <sup>36</sup>
	8.42 (shoulder)		
	9.82		9.83; <sup>33</sup> 9.9 <sup>36</sup>
	10.97		11.02; <sup>33</sup> 11.05 <sup>36</sup>
	11.55		11.45 <sup>36</sup>
	12.20		12.25 <sup>36</sup>
	12.62 (H <sub>2</sub> O)		
	13.02		13.00 <sup>36</sup>
	13.80		
	14.25		
	15.30		
	15.80		
	17.00		
	18.52		
	19.25		
	2,3-dihydro-7-azaindole	7.92	7.916 / 32a ( $\pi$ ) (HOMO)
9.42 (shoulder)		9.382 / 30a ( $n_N$ )	
9.81		9.823 / 31a ( $\pi$ )	
10.56		10.540 / 29a ( $\pi$ )	
12.15		12.303 / 28a ( $\sigma$ )	
		13.035 / 27a ( $\sigma$ )	
13.23 (broad)		13.157 / 26a ( $\sigma$ )	
		13.423 / 25a ( $\sigma$ )	
		13.517 / 24a ( $\sigma$ )	
		13.581 / 23a ( $\sigma$ )	
14.58		14.576 / 22a ( $\sigma$ )	
15.30		15.388 / 21a ( $\sigma$ )	
15.63		15.945 / 20a ( $\sigma$ )	
16.98		16.345 / 19a ( $\sigma$ )	
19.20 (broad)	17.365 / 18a ( $\sigma$ )		
3-formylindole	8.15		
	8.30	8.354 / 6a'' ( $\pi$ ) (HOMO)	
	8.72	8.684 / 5a'' ( $\pi$ )	
	9.11 (shoulder)		
	9.23	9.339 / 32a' ( $n_O$ )	
	10.06	10.105 / 4a'' ( $\pi$ )	
	11.40	11.663 / 3a'' ( $\pi$ )	
	11.90	12.035 / 31a' ( $\sigma$ )	
	12.52	12.485 / 30a' ( $\sigma$ )	
	13.35	13.012 / 2a'' ( $\pi$ )	
		13.618 / 29a' ( $\sigma$ )	
	14.23	14.247 / 28a' ( $\sigma$ )	
		14.566 / 27a' ( $\sigma$ )	
		14.308 / 1a'' ( $\pi$ )	
		14.514 / 26a' ( $\sigma$ )	
		14.851 / 25a' ( $\sigma$ )	
15.50	15.775 / 24a' ( $\sigma$ )		
	16.342 / 23a' ( $\sigma$ )		
17.30 (broad)	17.637 / 22a' ( $\sigma$ )		
18.70	18.184 / 21a' ( $\sigma$ )		

<sup>a</sup>High resolution data for the first two low-energy bands of indole.

ionizations. This is reflected in computed pole strengths (PS) smaller than 0.85.<sup>44</sup> EP methods used in this work include the outer valence Green function (OVGF),<sup>43,45</sup> partial third order (P3),<sup>46</sup> and renormalized partial third order (P3+)<sup>47</sup> methods, as implemented in Gaussian 09 in combination with the cc-pVTZ basis set. As mentioned above, these methods are of sufficient quality to yield reasonable agreement with experimental IPs in instances where the breakdown of the orbital picture of ionization does not occur (as seen from the computed pole strengths and also reported in the [Supporting Information](#)<sup>42</sup>). Several studies demonstrate that, within the diagonal self-energy approximation, P3+ combined with a basis set of triple-zeta quality represents the best compromise between accuracy and computational efficiency (i.e., see ref 48).

Regarding the core region, the C 1s, N 1s, and O 1s photoelectron spectra were calculated at the DFT level<sup>49</sup> with the hybrid B3LYP<sup>50</sup> and the generalized gradient approximation (GGA) PW86x Perdew<sup>51</sup> exchange correlation (xc) potentials and with the Amsterdam density functional (ADF) program.<sup>52</sup> The IPs were computed with the delta Kohn–Sham ( $\Delta$ -KS) scheme, namely, as differences between the KS solution in the ground state and ionic state. This calculation allows full relaxation of the ionized core hole within a spin-polarized unrestricted scheme.

**Experimental Methods.** Indole, 7-AI, and 3-FI were purchased from Sigma-Aldrich in the form of crystalline powder with a minimum purity of 99% and used without further purification. The samples were introduced into the system via an effusive needle source at room temperature (I and 7-AI) or evaporated from a crucible at a temperature of 350 K (3-FI) and a background pressure of 1 to  $7 \times 10^{-7}$  mbar. During the experiment, the sample quality was periodically monitored by valence band photoemission and photoionization mass spectroscopy.

The photoemission spectra were collected at the Gas Phase Photoemission and Circular Polarization beamlines at Elettra (Trieste, Italy).<sup>53,54</sup> The high-resolution C 1s, N 1s, and O 1s XPS and valence band spectra were measured using a VG-220i hemispherical electron energy analyzer. The analyzer was mounted in the plane perpendicular to the photon propagation direction and at an angle of  $54.7^\circ$  with respect to the electric vector of the light. In this geometry, the axis of the analyzer is set at the pseudo magic angle, and so measurements are insensitive to the photoelectron asymmetry  $\beta$  parameter. A homemade ion time-of-flight (TOF) mass spectrometer was mounted in the same chamber, facing the analyzer. This spectrometer configuration was successfully used to study various biological molecules in previous experiments.<sup>55</sup> Recently, the performance of this instrument was improved by replacing the six-channel electron multiplier detector unit of the analyzer by a rectangular microchannel plate detector<sup>56</sup> coupled to a 2D time delay-line anode developed by the Elettra detector laboratory.<sup>57,58</sup> The current detector covers  $\sim 12\%$  of the selected pass energy in a single exposure and thus requires shorter acquisition times.

In addition, the previous rectangular electron multipliers with a 2 mm opening in the dispersive direction acted also as exit slits. The analyzer resolution for a given pass energy is now governed mainly by the analyzer entrance lens optics as the detector acts as a virtual exit slit, adjustable simply by choosing the region to be integrated. In practice, under standard operation conditions, the kinetic energy resolution for pass

energies above 10 eV was improved by  $\sim 25\%$  with respect to the previous detector. The replacement of the detector has significantly improved the performance of the setup in energy-analyzed PEPICO experiments, which will be the subject of a separate publication.<sup>59</sup>

Valence spectra of the indoles were recorded with an incident photon energy of 60 eV; they are consistent with those previously measured with a He I (21.2 eV) source<sup>32–36</sup> apart from differences in relative peak intensity due to cross-sectional differences at the two photon energies. The binding energy (BE) scale for valence band spectra was calibrated using H<sub>2</sub>O (from residual gas).<sup>60</sup> The C 1s, N 1s, and O 1s core photoemission spectra were taken at 385, 495, and 628 eV of photon energy, respectively, and they were calibrated according to the well-known references of CO<sub>2</sub>, BE (C(1s))<sup>-1</sup> at 297.7 eV, O(1s)<sup>-1</sup> at 541.3 eV,<sup>61,62</sup> and N<sub>2</sub> (N(1s))<sup>-1</sup> at 409.9 eV.<sup>63</sup> The spectra were measured with a total resolution (photons + analyzer) of 0.20, 0.32, 0.46, and 0.53 eV at photon energies of 60, 385, 495, and 628 eV, respectively. At these photon energies, the kinetic energy of the photoelectrons is such that the effects due to post collisional interaction (PCI) can be neglected in the data analysis. Additionally, VB spectra for indole were taken at higher resolution of 50 meV in order to determine more accurately ionization potentials of the first two low-energy bands associated with its highest occupied molecular orbitals (HOMO and HOMO-1). The changes in transmission of the analyzer over the relatively small range of the photoelectron spectra shown were neglected.

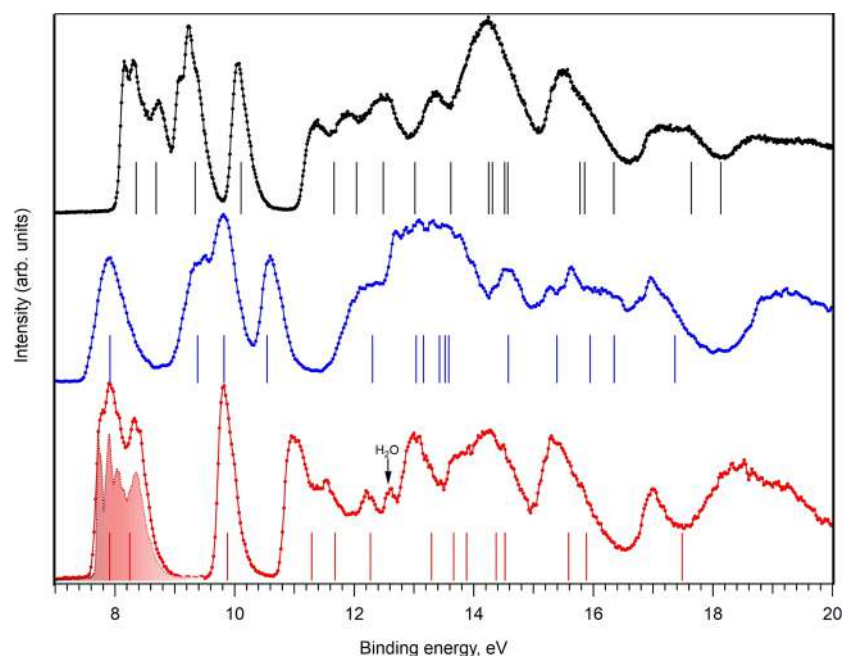
## RESULTS AND DISCUSSION

**Valence Photoemission Spectra.** For all three molecules, we performed OVGF, P3, and P3+ calculations in order to identify the most accurate EP method for reproducing the valence photoemission spectra. The complete list of calculated energies, together with the corresponding pole strengths, is given in the [Supporting Information](#).<sup>42</sup> The assignments of the observed features of the outer valence band based on the P3+ calculations are summarized in [Table 1](#). Additionally, the present data have been compared with previously reported experimental values obtained using He I radiation.<sup>33,35,36</sup>

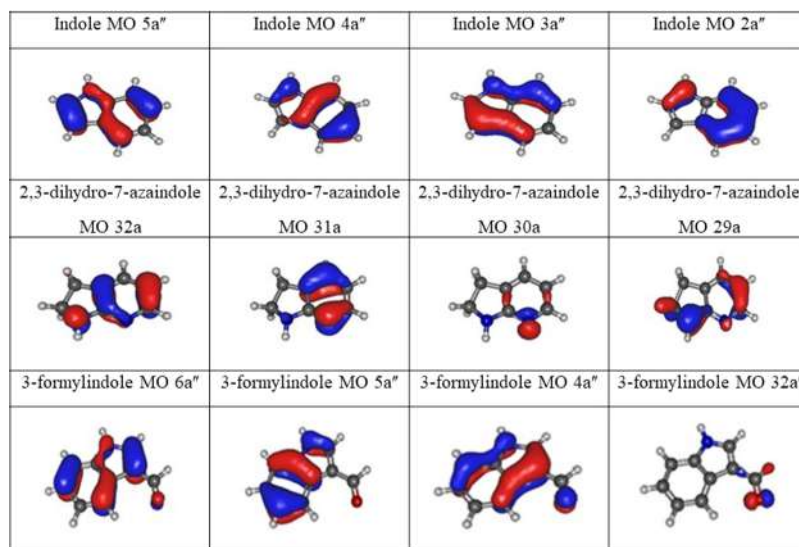
Before describing our photoelectron spectra, we note that the three studied heterocycles have chemical differences (see [Figure 1](#)). For example, indole is a benzo-indole, a planar compound, and 7-azaindole is a non-planar compound due to the sp<sup>3</sup> hybridization of C2 and C3 atoms in the pentagonal ring. The pair differs in both the hexagon and pentagonal rings. Indole and 3-formylindole are both benzo-indoles that differ by the 3-formyl group in the pentagonal ring, but the latter has both *trans*- and *cis*-conformers, which are detectable.<sup>30</sup> Hence, the difference in their chemical structure will influence the shape and peak positions of the measured PES spectra.

The photoelectron spectrum of indole exhibits a low-energy band at 7.90 eV, which is associated with its highest occupied molecular orbital (HOMO) of  $\pi$  symmetry ( $5a''$ ). The next three bands, appearing at 8.32, 9.82, and 10.97 eV, correspond to the  $4a''$ ,  $3a''$ , and  $2a''$  MOs, respectively. These MOs also have  $\pi$  character and are delocalized on the indole ring (see [Figure 3](#)). Our assignment is consistent with the available valence photoemission spectra for indole,<sup>33,36</sup> moreover, the ionization energies of the lowest MOs are well reproduced by the present P3+ calculations (see [Table 1](#)). The first three peaks are mostly found at the predicted energies, while for the fourth ionization ( $2a''$ ), we find a discrepancy of 0.3 eV with





**Figure 2.** Dotted lines: valence-band photoelectron spectra of indole (bottom curve, red), 2,3-dihydro-7-azaindole (center curve, blue), and 3-formylindole (top curve, black) at a photon energy of 60 eV. Shaded spectrum: higher-resolution valence band spectrum of indole. Bars: calculated BEs for each compound using the P3+ cc-pVTZ model.



**Figure 3.** Charge density maps of the outermost molecular orbitals of indole, 2,3-dihydro-7-azaindole, and 3-formylindole.

respect to the experimental value. The computed low PS value (0.815) for this MO indicates however a likely breakdown of the quasi-particle picture of ionization (see Table S1 of the Supporting Information).<sup>42</sup>

The higher resolution VB spectrum of indole (see Figure 2, shaded spectrum) consists of several bands at low IPs accompanied by vibrational splitting of  $\pi_1$  (7.72, 7.76, and 7.90 eV) and  $\pi_2$  (8.05 and 8.35 eV) orbitals. This spectrum is in good agreement with that of DomelSmith and co-workers<sup>33</sup> with BEs of 7.75, 7.92, 8.08, and 8.37 eV. The indole structure consists of fused pyrrole and benzene rings (see Figure 1). For comparison, pyrrole has its first two valence orbitals at 8.2 and 9.2 eV, while those of benzene are at 9.4 and 11.8 eV.<sup>64</sup>

The substitution of one carbon atom in the benzene ring of indole by nitrogen (see Figure 1b) forms 2,3-dihydro-7-

azaindole (7-AI), which has a different photoemission spectrum. Similar to indole, the first broad band located at 7.92 eV was assigned to the HOMO. The vertical ionization potential found here for 7-AI is slightly lower ( $\sim 0.2$  eV) than that published by Fuke et al.<sup>35</sup> The second feature with maxima at 9.42 and 9.81 eV is close to the assignment found for the lowest IP of pyridine, 9.6 eV.<sup>60,65</sup> The first region of the VB ionization in pyridine has been attributed to both  $\pi$ - and  $n$ -orbital ionization, and photoemission spectra support the view that there are two overlapping bands. The comparison of our experimental and computational data shows that the replacement of one carbon atom in indole by nitrogen does not result in significant changes in the energy levels of the corresponding HOMOs. According to our calculations, the features in the valence photoemission spectrum of the 7-AI

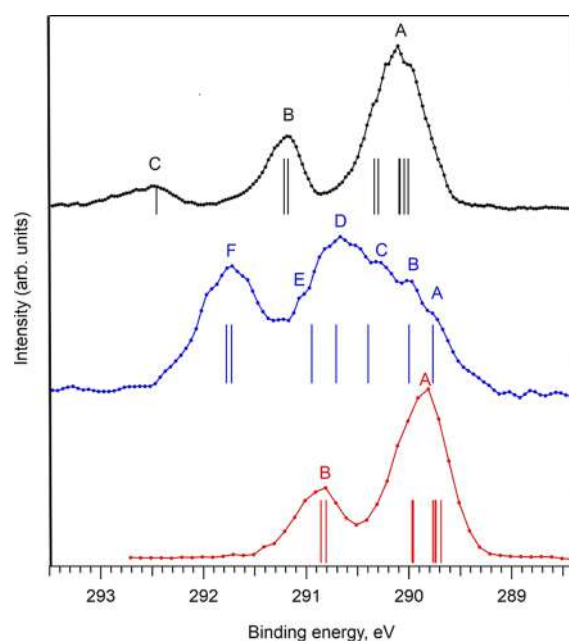
molecule at 9.81 and 10.56 eV are assigned to the MO with  $\pi$  symmetry, while the shoulder at 9.42 eV is due to the nitrogen heteroatom lone pair (see Figure 3). In the case of 7-AI, EP methods produce a different ordering of the cationic states with respect to that of the KT. In particular, ionization energies related to the  $\pi$  (31a) and  $n$  (30a) MOs are reversed. This shows that, for the cases of nitrogen-containing heterocycles, EP corrections to Koopmans' results are essential to produce correct assignments of photoelectron spectra.<sup>66,67</sup> In fact, correlation and relaxation effects are generally much more relevant for hole states with dominant contributions from nonbonding, nitrogen-centered functions than for delocalized  $\pi$  MOs.

The first photoelectron peak of 3-FI is split into peaks at 8.15 and 8.30 eV and corresponds to the HOMO ionization. From the computed relative energies of 1H-*trans* and 1H-*cis* isomers of 3-FI, the relative Boltzmann populations of these two forms in the gas phase at 350 K were estimated as  $\sim 85\%$  1H-*trans* and  $\sim 15\%$  1H-*cis*.<sup>30</sup> Hence, the splitting of the first maximum could be due to the presence of different conformers.

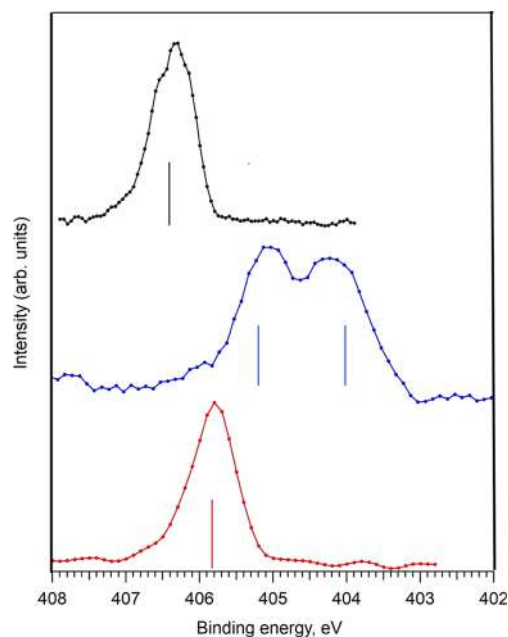
The next bands at 8.72 and 10.06 eV are assigned to MOs (HOMO-1 and HOMO-2) of  $\pi$  symmetry. The asymmetric feature at 9.23 eV is due to the ionization of the 32a' MO, which is mostly located on the oxygen heteroatom lone pair (see Figure 3). This assignment is in good agreement with already published valence band photoemission data for tryptophol, which also contains an indole ring but differs in having a flexible ethanolic side chain.<sup>40</sup>

As for 3-FI, the KT results are quantitatively inadequate, and they fail to predict the right order of final states. The correct ordering is retained by using EP methods, which consider both electron correlation and orbital relaxation effects, particularly relevant for final states with lone pair holes. However, while, in the case of 7-AI, both OVGf and P3+ methods correctly reproduce the band related to the ionization from the lone-pair, for the 3-FI molecule, P3+ results agree with the experiment better than those from OVGf and P3 (see the Supporting Information, Tables S1–S3).<sup>42</sup> In particular, OVGf overestimates the energy related to the lone-pair band by 0.52 eV, while both P3 and P3+ properly predict that energy, providing results in good agreement with the experimental data (see Table S3, Supporting Information).<sup>42</sup> Moreover, a comparison of the computed PS for the valence ionization of 7-AI and 3-FI reveals that the latter is a more challenging system for all diagonal EP approximations: with the exception of the outermost HOMO, HOMO-1, and HOMO-2 ionizations of  $\pi$ -type MOs, for all other  $\pi$  MOs, the low values of the corresponding PS indicates the importance of orbital relaxation effects. In fact, significant relaxation effects are anticipated for all ionization energies above 18 eV (see Table S3, Supporting Information).<sup>42</sup>

**Core Level Photoemission.** The C 1s, N 1s, and O 1s XPS experimental and theoretical photoemission spectra of the samples are shown in Figures 4 and 5 and Figure S4 of the Supporting Information (O 1s XPS)<sup>42</sup> and summarized in Table 2. Experimentally derived BEs are compared with theoretical BEs computed by using the hybrid B3LYP xc potential. The assignment of the spectral features is rationalized considering that core level BEs are very sensitive to the chemical environment of the ionized atomic site in terms of both electron density and electronic relaxation. Starting with I, the carbon core-level spectrum (see Figure 4)



**Figure 4.** C 1s photoemission spectra of indole (bottom curve, red), 2,3-dihydro-7-azaindole (center curve, blue) and 3-formylindole (top curve, black). Dotted lines: experimental data. Bars: theoretical data computed by using the hybrid B3LYP xc potential. Calculated BEs were shifted by +0.2, +0.15, and +0.1 eV for indole, 2,3-dihydro-7-azaindole, and 3-formylindole, respectively.



**Figure 5.** N 1s photoemission spectra of indole (bottom curve, red), 2,3-dihydro-7-azaindole (center curve, blue) and 3-formylindole (top curve, black). Dotted lines: experimental data. Bars: theoretical data computed by using the hybrid B3LYPxc potential. Calculated BEs were shifted by +0.35 eV for all three indoles.

shows two distinct features (labeled A and B) centered at 289.89 and 290.86 eV with an integrated intensity ratio of 6:2. Indole is formed by the fusion of a pyrrole ring to the benzene ring at the C9=C8 position (see Figure 1a). Our calculations show that all the carbon atoms not bonded to nitrogen contribute to the strongest peak A of the photoemission

**Table 2.** C, N, and O 1s Experimental and Calculated BEs for Indole, 2,3-Dihydro-7-azaindole, and 3-Formylindole

molecule/ core level	experimental BEs (eV) ± 0.1 eV/feature	calculated BEs (eV) ± 0.1 eV, DFT: B3LYPxc/PW86x/Assignment
indole		
C 1s	289.89/A	289.49/289.86/C5
		289.54/289.93/C3
		289.55/289.91/C6
		289.57/289.93/C4
		289.76/290.16/C9
		289.77/290.15/C7
		290.86/B
N 1s	405.82	290.61/290.94/C2
		290.66/291.01/C8
2,3-dihydro-7-azaindole		
C 1s	289.77/A	289.62/289.99/C5
		290.00/B
		289.85/290.26/C9
		290.27/C
		290.56/290.84/C6
		290.67/D
		291.02/E
N 1s	405.03	291.58/291.95/C2
		291.63/291.87/C8
		404.23
		403.56/404.17/N7
		404.69/405.35/N1
3-formylindole		
C 1s	290.25/A	289.91/290.34/C5
		289.95/290.37/C6
		289.99/290.42/C4
		290.00/290.45/C3
		290.20/290.63/C7
		290.24/290.67/C9
		291.33/B
N 1s	406.36	291.08/291.41/C2
		291.12/291.51/C8
O 1s	536.84	292.36/292.45/C10
		406.04/406.71/N1

spectrum, while peak B is due to the C2 and C8 carbons bonded to the N atom, as summarized in Table 2. Their higher BE values are due to their higher electrostatic potential with respect to the other C sites. The observed C 1s BEs are in good agreement with previously published data for 3-methylindole,<sup>39</sup> tryptophol,<sup>40</sup> tryptamine,<sup>40</sup> benzene, and pyrrole.<sup>64</sup> The BE of the benzene C 1s level was found to be 290.20 eV<sup>64</sup> and it is about 0.3 eV higher than the measured energy of peak A. The small shift in BE can be attributed to the final-state screening, because indole is larger and more polarizable than the benzene ring.

The C 1s spectra for I and 3-FI are, as expected, quite similar, given that the molecules differ by one hydrogen atom being replaced by a formyl group (see Figure 1). The area ratios of peaks A:B:C in the spectrum of 3-FI are 6:2:1. We notice that the position of peaks A and B is shifted to higher BEs (by about 0.30 eV) on going from I to 3-FI, due to the presence of the formyl group in 3-FI. In the C 1s XPS spectrum of 3-FI, the broad peak C at 292.65 eV is assigned to the ionization of the carbon atom bonded to the electronegative oxygen atom which causes an increase of the electrostatic potential at the carbon nucleus (see Figure 4). As in the case of indole, tryptophol, tryptamine and 3-methylindole, the most pronounced feature A, centered at 290.25 eV in the carbon core level spectrum of 3-FI, is due to

the aromatic ring carbons<sup>39,40,64</sup> which are not directly bonded to N, while feature B is associated with ionization from C8 and C2 (see Table 2).

The experimental full widths at half maximum of the most intense features A (see Figure 4) in the C 1s spectra of indole and 3-formylindole, are about 0.89 and 0.84 eV, respectively. Assuming Gaussian peak shapes and considering the total experimental resolution of 0.32 eV, the observed widths derived are 0.79 and 0.74 eV. The number of closely spaced electronic states and their vibrational envelopes hamper a better definition of the structure, even for improved experimental resolution.

7-AI contains seven carbon atoms, and wide asymmetric peaks are seen in its C 1s photoemission spectrum (see Figure 4). The low BE feature of the C 1s XPS spectrum of 7AI is noticeably broader than for the other two indoles, and shows five well resolved shoulders (A-E) lying in the range 289–291 eV (see Figure 4). The distinct shifts of the carbon 1s BEs of 2,3-dihydro-7-azaindole are due to the different chemical environment and could be characterized as a superposition of carbon atoms from the pyridine and pyrrole rings. The experimental C 1s BEs for pyridine were found to be 290.9 eV, 290.6 eV and 290.2 eV for (C2 = C6), (C4), and (C3 = C5) core holes, respectively.<sup>68,69</sup> For pyrrole, two carbon core-electron BEs at 290.8 eV (C2 = C5) and 289.8 eV (C3 = C4) have been published by Gelius et al.<sup>64</sup> It is interesting to note the manner in which C 1s BEs values for the 7-AI molecule interchange with respect to the position of the carbon atom in a ring system such as pyridine, in which the lowest carbon ionization values were obtained for the C9 and C5 atoms located as nearest-neighbors to the (N7) atom (see Table 2). The same trend was observed for various carbon–nitrogen molecules by Snis et al.<sup>68</sup> The highest BE feature F at 291.72 eV has been assigned as for indole and 3-formylindole, and is due to the two carbon atoms (C2 and C8) which are nearest-neighbors of nitrogen (N1) in the pyrrole ring (see Figure 1). The values are about 0.86 and 0.39 eV higher than the energy of the same carbons in the indole and 3-FI molecules, respectively (see Table 2). This can be explained by considering the electronegativity of the neighbouring atoms. For the case of 7-AI, the electronegative N-atom directly bonded to C8 produces a decrease of the electron density around the C-atom and therefore an increase of the binding energy. Such an effect is reduced in the case of 3-FI, as the electronegative oxygen atom is not directly bonded to the considered C atoms.

The N 1s photoemission spectra of indole and 3-FI present broad single features at 405.82 and 406.36 eV, respectively (see Figure 5). In both cases the peak derives from the same chemical environment and is due to the nitrogen atom (N1) in the pyrrole ring. The experimental N 1s BE for pyrrole was determined to be 406.1 eV,<sup>64</sup> and it is approximately 0.3 eV higher than the value obtained for indole and lower than those measured for 3-methylindole. The BE of the indole nitrogen matches well the already published values for 3-methylindole (405.7 eV),<sup>39</sup> tryptophol, and tryptamine (405.73 eV).<sup>40</sup>

The N 1s spectra of indole and 3-formylindole both show asymmetric peaks, a shape which has been attributed to Franck–Condon effects.<sup>39</sup> The experimental fitted widths are about 1.03 and 0.94 eV for indole and 3-FI, respectively. Taking into account the experimental resolution, this implies intrinsic widths of 0.82 and 0.73 eV, similar to the values obtained from the C 1s spectra.

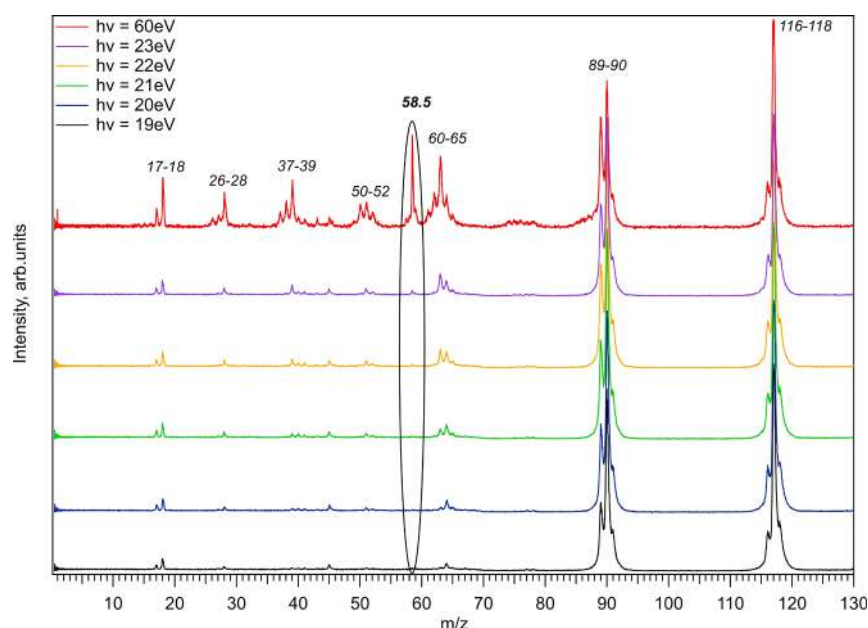


Figure 6. Mass spectra of indole as a function of photon energy.

In 7-AI, two nitrogen atoms are present, and two peaks are observed in the experimental spectrum at 404.23 and 405.03 eV (see Figure 5). The difference in energy between these features is 0.8 eV, slightly lower than that predicted by theoretical calculation ( $\sim 1.15$  eV) (see Table 2). The peak at higher BE corresponds to the N1 atom in the pyrrole ring, which has the same pattern as that found for indole, 3-formylindole, 3-methylindole, tryptophol, and tryptamine measured under similar experimental conditions. The lower-energy peak at 404.23 eV is assigned to the N atom of the pyridine ring. Their relative positions can be understood by simple arguments based on the different electron shielding effects on the two N positions. The experimental position of the N 1s peak corresponding to the ionization of the N1 atom for 7-AI is approximately 0.8 and 1.3 eV lower with respect to the single features observed for I and 3-FI, respectively. The closest calculated N 1s BEs found for pyridine are 404.69 eV<sup>68</sup> and 404.76 eV,<sup>69</sup> approximately 0.5 eV higher than the N7 atom energy in 7-AI. Also, experimental values of nitrogen core-electron BEs of pyridine show a shift, which is approximately 0.7 eV higher<sup>69</sup> with respect to the value measured in the present work. It is well known that, in the case of organic heterocycles, the ionization energy decreases with the increasing size of the system.<sup>68</sup>

The computed and measured O 1s core level spectra of 3-formylindole are displayed in Figure S4 (see the Supporting Information),<sup>42</sup> while the corresponding BEs are summarized in Table 2. There was no effect of the molecular conformation observed in the experimental O 1s spectrum of 3-FI. The intense peak at 536.84 eV was assigned to the oxygen atom of the formyl group in the molecule. In the case of 3-FI the O 1s BE is lower than for tryptophan and tryptophol, which contain the indole structure and an additional carboxyl or hydroxyl group.<sup>39,40</sup> Based on its energy, we attribute the wide asymmetric peak at 539.80 eV to water present in the sample and/or the experimental chamber.<sup>70</sup>

**Doubly Charged Ions of Indole.** Mass spectrometry has provided a large amount of data for doubly charged ions (dications).<sup>71–73</sup> The minimum energy required to form the

doubly charged ion is denoted variously in the literature as the double ionization potential (DIP), double ionization energy (DIE), or appearance potential (AP). In this section, we report data for the double ionization of indole.

The parent ion ( $M^+$ ) of indole  $m/z = 117$  dominates the mass spectrum at all photon energies (see Figure 6). It has been shown that, for aromatic heterocyclic fused-ring compounds containing more than one nitrogen atom, in most cases, the dominant fragmentation reaction of  $M^+$  involves expulsion of  $\text{HCN}^\circ$  (or  $\text{HNC}^\circ$ ) neutral species ( $m = 27$ ).<sup>73</sup> Consistent with this, at all the photon energies used, the next most intense pair of masses after the parent ion are at  $m/z = 90$  and  $m/z = 89$  corresponding to  $\text{C}_7\text{H}_6^+$  and  $\text{C}_7\text{H}_5^+$  fragments. The very weak signal at  $m/z = 58.5$  in the spectrum measured at  $h\nu = 22$  eV is due to the formation of the dication. As the photon energy increases, the signal of the doubly charged ion becomes more pronounced (see Figure 6,  $h\nu = 60$  eV). In general, the probability for double ionization is always lower than that of the single ionization. Note that our photofragmentation spectra of indole are in good agreement with those measured by electron impact at 70 eV.<sup>65</sup> Additionally, fragmentation reactions of both metastable and collisionally activated dications, formed by electron impact ionization of 25 polycyclic aromatic compounds have been investigated in detail by Perreault et al.<sup>73</sup> Their results have been interpreted on the basis of mechanistic models based on the competition between charge-separation and neutral-expulsion reactions and on variations in the diradical character with increasing molecular size and nitrogen content of these molecular dications.

To estimate the appearance potential of the doubly charged ion of indole, we chose Auger electron spectroscopy<sup>74</sup>, as the final states of this process are doubly charged. The subtraction of the Auger electron energy from the energy of the initial hole state gives the energy of the state of the ion.<sup>74</sup> The C-Auger spectrum of I measured here shows the first peak centered at a BE of approximately 21.8 eV, which corresponds to its lowest DIE (see Figure 7). This value is consistent with the position of the first wide band presented in the N-Auger spectrum of

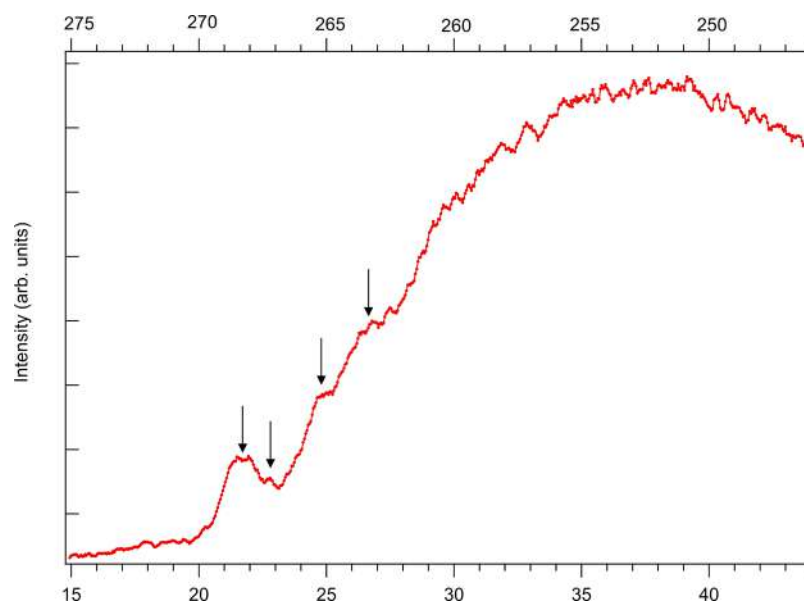


Figure 7. C-Auger spectrum of indole measured at a photon energy of 350 eV.

indole (see the Supporting Information, Figure S5) as well as with our TOF spectra as a function of photon energies (Figure 6). It is not surprising that the present C- and N-Auger spectra of indole have broad structures. This broadening includes the finite line width of the XPS, the vibrational distributions of the core-hole states and the final states, the core-hole lifetime, and also the analyzer resolution. Note that the energy scales of the C- and N-Auger spectra in Figure 7 and Figure S5 (Supporting Information) have been calibrated with respect to CO<sub>2</sub><sup>75</sup> and N<sub>2</sub><sup>76</sup> and the experimentally found BE values 289.89 and 405.82 eV for carbon and nitrogen core-level photoemission spectra of indole, respectively (see Table 2).

In order to understand which orbitals/states of indole are involved in the formation of doubly charged ions, we have performed calculations at the CASSCF/MRCI (complete active space self-consistent field/multi-reference internally contracted configuration interaction) level of theory<sup>77</sup> with the ANO-VT-TZ basis, and these data are summarized in Table S4 of the Supporting Information. All calculations were performed by using the MOLPRO package.<sup>78</sup> In detail, the CASSCF energies were obtained by averaging over the four lowest singlet states and one triplet state with an active space including eight active electrons for the neutral state and six active electrons for the doubly ionized states in six active orbitals (see the Supporting Information, Figures S6 and S7). Although all the calculated states of the doubly charged molecule have mixed configurations, they are mainly dominated by the contribution of one of the reference CI configurations (see Supporting Information, Table S4).<sup>42</sup>

As mentioned above, there are few singlets and one triplet state relatively close in energy (see the Supporting Information, Table S4 and Figures S6 and S7) that could participate in the dication formation. Our calculations show that the first two states ( $1^1A'$  and  $2^1A'$ ) are separated by 0.5 eV, while the third state ( $3^1A'$ ) is 1.5 eV higher. As we do not calculate the intensities, the peak shape is difficult to predict. The calculations also show that the next state ( $4^1A'$ ) is much higher in energy, so it would seem that we see the decay to the lowest states for C and for N (see the Supporting Information, Table S4).<sup>42</sup> Note that the lowest triplet state ( $1^3A'$ ) is

energetically near (0.18 eV) the lowest singlet state ( $1^1A'$ ) (see the Supporting Information, Table S4). One can assign “dominant configurations” to the lowest states, and we consider holes in the HOMO and HOMO-1. According to Figure 3, these molecular orbitals ( $5a''$  and  $4a''$ ) extend over the whole molecule, including the nitrogen atom. Hence, the first wide peak in the N-Auger spectrum (see the Supporting Information, Figure S5) is most likely due to the lowest dication state of indole. The same is true for the C-Auger spectrum (see Figure 7), but things are more complicated here since the molecule has many carbon atoms. The C 1s photoemission spectra of indole show two peaks: one having contributions of C2 and C8 (bound to the nitrogen) at 290.86 eV and the other at 289.89 eV with contributions from all other carbons (see Table 2 and Figure 4). For the lowest dicationic state, this implies that at least two peaks separated by  $\sim 1$  eV should be observed in the C-Auger spectrum of indole. In Figure 7, one can see a small shoulder at 22.8 eV, which may tentatively be assigned to this process. From Figure 3, it appears that the two outermost orbitals (MO  $5a''$  and MO  $4a''$ ) extend over C2 and C8 atoms but also over most of the other carbon atoms complicating the final interpretation. Therefore, one could assume that the first observed structure in the C-Auger spectrum is at an energy consistent with a decay from some of the (not nitrogen neighbor) ring carbons to the lowest state of the doubly charged ion. Additionally, our mass spectra are also consistent with the energy obtained from the Auger spectra, showing that the dication is stable on the timescale of our experiment (hundreds of nanoseconds).

Other weak shoulders around 24.84 and 26.78 eV have been observed as well in the C-Auger spectrum of indole. Since no prior experimental and theoretical studies of the double ionization of indoles exist in the literature, we can only compare our data with that of other aromatic compounds (see Table 3).<sup>72,79,80</sup> Dication appearance potentials for aromatic molecules containing a double ring structure (such as indole, quinoline, and azulene) are lower than those having a single-ring structure (benzene, pyridine and pyrimidine). It is certainly possible to predict double ionization appearance potentials from single ionization energies using some empirical

**Table 3. Double Ionization Appearance Potentials of Some Aromatic Molecules**

molecule/published data	first peak (DIE) (eV)	other peaks/shoulders
indole (present work)	21.8 ± 0.2	22.8, 24.84, 26.78
pyridine <sup>72</sup>	25.1	from 25.6 to 32.0
pyrimidine <sup>79</sup>	25.4	from 27.61 to 32.16
quinoline <sup>70</sup>	22.4	23.2, 26
benzene <sup>80</sup>	24.65	from 26 to 31
azulene <sup>80</sup>	20.2	21.3, 22.3

equations, provided that one knows the distance between the final charges of the molecule.<sup>72,81</sup> However, this information is often lacking; therefore, this empirical approach is presently of little practical use.

## CONCLUSIONS

The core and valence photoemission spectra of three biologically important compounds, namely, indole, 2,3-dihydro-7-azaindole, and 3-formylindole, have been studied in the gas phase using X-ray photoelectron spectroscopy (XPS) and quantum-chemistry calculations. Comprehensive information about the electronic structure and, in particular, the occupied molecular orbitals, has been acquired. The high-resolution spectrum for the inner valence region of indole has been reported, and several bands due to the vibrational splitting of its HOMO and HOMO-1 have been defined. The appearance potential of the doubly charged ion of indole has been determined to be 21.8 ± 0.2 eV by using Auger electron spectroscopy. The carbon, nitrogen, and oxygen 1s photoemission spectra for all three investigated molecules have been assigned. Additionally, calculations presented in this paper are in good agreement with the relative BEs of the core level features observed in the experimental photoemission results. By comparing our data with previously published results, we conclude that the rotation of the formyl group attached in 3-FI to the C3 atom was weak and did not show detectable chemical shifts in the core-level spectra of this molecule.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpca.0c02719>.

Complete list of calculated energies together with the corresponding pole strengths for indole, 2,3-dihydro-7-azaindole, and 3-formylindole; full plot HF/cc-pVTZ orbitals for three nitrogen-containing heterocycles; O 1s photoemission spectrum of 3-formylindole and N-Auger spectrum of indole; and calculated double ionization energy values for the indole molecule (PDF)

## AUTHOR INFORMATION

### Corresponding Author

**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

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

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

**Giuseppe Cautero** – *Elettra-Sincrotrone Trieste S.C.p.A., 34149 Trieste, Italy*

**Carlo Dri** – *Elettra-Sincrotrone Trieste S.C.p.A., 34149 Trieste, Italy; IOM-CNR Laboratorio TASC, 34149 Trieste, Italy*; [orcid.org/0000-0001-9040-5746](https://orcid.org/0000-0001-9040-5746)

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

**Kevin C. Prince** – *Elettra-Sincrotrone Trieste S.C.p.A., 34149 Trieste, Italy; Centre for Translational Atomaterials, Swinburne University of Technology, Melbourne 3122, Australia*

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

**Rudi Sergio** – *Elettra-Sincrotrone Trieste S.C.p.A., 34149 Trieste, Italy*

**Luigi Stebel** – *Elettra-Sincrotrone Trieste S.C.p.A., 34149 Trieste, Italy*

**Michele Devetta** – *Istituto di Fotonica e Nanotecnologie, CNR, 20133 Milano, Italy*

**Davide Faccialà** – *Istituto di Fotonica e Nanotecnologie, CNR, 20133 Milano, Italy*

**Caterina Vozi** – *Istituto di Fotonica e Nanotecnologie, CNR, 20133 Milano, Italy*

**Lorenzo Avaldi** – *CNR - Istituto di Struttura della Materia (CNR-ISM), 00133 Rome, Italy*

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

**Mattea Carmen Castrovilli** – *CNR - Istituto di Struttura della Materia (CNR-ISM), 00133 Rome, Italy*; [orcid.org/0000-0002-7909-5115](https://orcid.org/0000-0002-7909-5115)

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

**Marcello Coreno** – *CNR - Istituto di Struttura Della Materia (CNR-ISM), 34149 Trieste, Italy*

**Fabio Zuccaro** – *CNR - Istituto di Struttura Della Materia (CNR-ISM), 34149 Trieste, Italy*

**Elisa Bernes** – *Dipartimento di Scienze Chimiche, Università degli Studi di Trieste, 34127 Trieste, Italy*

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

**Daniele Toffoli** – *Dipartimento di Scienze Chimiche, Università degli Studi di Trieste, 34127 Trieste, Italy*; [orcid.org/0000-0002-8225-6119](https://orcid.org/0000-0002-8225-6119)

**Aurora Ponzi** – *Institut Ruđer Bošković, 10000 Zagreb, Croatia*; [orcid.org/0000-0001-9095-4366](https://orcid.org/0000-0001-9095-4366)

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpca.0c02719>

## 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 has been supported by the Croatian Science Foundation under grant no. HRZZ IP-2016-06-1142. This work has been 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 has been supported by Finanziamento per Ricerca di Ateneo, FRA 2015 and FRA 2016, of the Università degli Studi di Trieste.

## REFERENCES

- (1) Sundberg, R. J. *The Chemistry of Indoles*; Academic Press: New York, 1996.
- (2) Meredith, P.; Sarna, T. The Physical and Chemical Properties of Eumelanin. *Pigm. Cell Res.* **2006**, *19*, 572–594.
- (3) Fuke, K.; Yoshiuchi, H.; Kaya, K. Electronic Spectra and Tautomerism of Hydrogen-Bonded Complexes of 7-Azaindole in a Supersonic Jet. *J. Phys. Chem.* **1984**, *88*, 5840–5844.
- (4) Sakota, K.; Okabe, C.; Nishi, N.; Sekiya, H. Excited-State Double-Proton Transfer in the 7-Azaindole Dimer in the Gas Phase. 3. Reaction Mechanism Studied by Picosecond Time-Resolved REMPI Spectroscopy. *J. Phys. Chem. A* **2005**, *109*, 5245–5247.
- (5) Fuke, K.; Ishikawa, H. Dynamics of Proton Transfer Reactions of Model Base Pairs in the Ground and Excited States: Revisited. *Chem. Phys. Lett.* **2015**, *623*, 117–129.
- (6) Zhang, L. S.; Davies, S. S. Microbial Metabolism of Dietary Components to Bioactive Metabolites: Opportunities for New Therapeutic Interventions. *Genome Med.* **2016**, *8*, 46.
- (7) 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.
- (8) Ghita, M.; Arrigan, D. W. Electrochemical Overoxidation of Polyindole and Its Cation-Permeable Behavior. *Electroanalysis* **2004**, *16*, 979–987.
- (9) Roman, G.; Pappas, A. C.; Kovala-Demertzi, D.; Prodromidis, M. I. Preparation of a 2-(4-Fluorophenyl)Indole-Modified Xerogel and Its Use for the Fabrication of Screen-Printed Electrodes for the Electrocatalytic Determination of Sulfide. *Anal. Chim. Acta* **2004**, *523*, 201–207.
- (10) Udum, Y. A.; Dündükcü, M.; Köleli, F. Electrochemical Polymerization and Spectroscopic Investigation of 2-Methylindole. *React. Funct. Polym.* **2008**, *68*, 861–867.
- (11) Talbi, H.; Monard, G.; Loos, M.; Billaud, D. Theoretical Study of Indole Polymerization. *J. Mol. Struct.: THEOCHEM* **1998**, *434*, 129–134.
- (12) Philips, L. A.; Levy, D. H. The Rotationally Resolved Electronic Spectrum of Indole in the Gas Phase. *J. Chem. Phys.* **1986**, *85*, 1327–1332.
- (13) Berden, G.; Meerts, W. L.; Jalviste, E. Rotationally Resolved Ultraviolet Spectroscopy of Indole, Indazole, and Benzimidazole: Inertial Axis Reorientation in the  $S_1(^1L_b) \leftarrow S_0$  Transitions. *J. Chem. Phys.* **1995**, *103*, 9596–9606.
- (14) Brand, C.; Küpper, J.; Pratt, D. W.; Meerts, W. L.; Krügler, 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.
- (15) 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.
- (16) Korter, T. M.; Pratt, D. W.; Küpper, J. Indole–H<sub>2</sub>O in the Gas Phase. Structures, Barriers to Internal Motion, and  $S_1 \leftarrow S_0$  Transition Moment Orientation. Solvent Reorganization in the Electronically Excited State. *J. Phys. Chem. A* **1998**, *102*, 7211–7216.
- (17) Kang, C.; Korter, T. M.; Pratt, D. W. Experimental Measurement of the Induced Dipole Moment of an Isolated Molecule in Its Ground and Electronically Excited States: Indole and Indole–H<sub>2</sub>O. *J. Chem. Phys.* **2005**, *122*, 174301.
- (18) Nesvadba, R.; Studecký, T.; Uhlíková, T.; Urban, Š. Microwave Spectrum and Molecular Constants of Indole. *J. Mol. Spectrosc.* **2017**, *339*, 6–11.
- (19) Short, K. W.; Callis, P. R. Evidence of Pure  $^1L_b$  Fluorescence from Redshifted Indole–Polar Solvent Complexes in a Supersonic Jet. *J. Chem. Phys.* **1998**, *108*, 10189.
- (20) 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.
- (21) 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*, 194307.
- (22) 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*, 25197–25209.
- (23) 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.
- (24) Oliver, T. A. A.; Kingz, G. A.; Ashfold, M. N. R. Position Matters: Competing O–H and N–H Photodissociation Pathways in Hydroxy- and Methoxy-Substituted Indoles. *Phys. Chem. Chem. Phys.* **2011**, *13*, 14646–14662.
- (25) Callis, P. R.; Vivian, J. T.; Slater, L. S. Ab Initio Calculations of Vibrational Spectra for Indole. *Chem. Phys. Lett.* **1995**, *244*, 53–58.
- (26) Sundaraganesan, N.; Umamaheswari, H.; Joshua, B. D.; Meganathan, C.; Ramalingam, M. Molecular Structure and Vibrational Spectra of Indole and Ab Initio Hartree-Fock Calculations. *J. Mol. Struct., THEOCHEM* **2008**, *850*, 84–93.
- (27) Catalán, J.; de Paz, J. L. G. The Molecular Geometry of Indole. *J. Mol. Struct. THEOCHEM* **1997**, *401*, 189–192.
- (28) Smith, B. J.; Liu, R. A Theoretical Investigation of Indole Tautomers. *J. Mol. Struct. THEOCHEM* **1999**, *491*, 211–222.
- (29) Caminati, W.; Di Bernardo, S. Microwave Spectrum and Amino Hydrogen Location in Indole. *J. Mol. Struct.* **1990**, *240*, 253–262.
- (30) 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*, 194304.
- (31) Nowak, M. 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.
- (32) Eland, J. H. D. Photoelectron Spectra of Conjugated Hydrocarbons and Heteromolecules. *Int. J. Mass Spectrom. Ion Phys.* **1969**, *2*, 471–484.
- (33) 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.
- (34) Kovač, B.; Klasinc, L.; Stanovnik, B.; Tišler, M. Photoelectron Spectroscopy of Heterocycles. Azaindenes and Azaindolizines. *J. Heterocycl. Chem.* **1980**, *17*, 689–694.
- (35) Fuke, K.; Yoshiuchi, H.; Kaya, K.; Achiba, Y.; Sato, K.; Kimura, K. Multiphoton Ionization Photoelectron Spectroscopy and Two-Color Multiphoton Ionization Threshold Spectroscopy on the Hydrogen Bonded Phenol and 7-Azaindole in a Supersonic Jet. *Chem. Phys. Lett.* **1984**, *108*, 179–184.
- (36) 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.

- (37) Kierspel, T.; Bomme, C.; Di Fraia, M.; Wiese, J.; Anielski, D.; Bari, S.; Boll, R.; Erk, B. M.; Kienitz, J. S.; Müller, N. L. M.; et al. *Phys. Chem. Chem. Phys.* **2018**, *20*, 20205–20216.
- (38) 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.
- (39) 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.
- (40) Maris, A.; Melandri, S.; Evangelisti, L.; Caminati, W.; Giuliano, B. M.; Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; Prince, K. C. Soft X-Ray Photoemission Spectroscopy of Selected Neurotransmitters in the Gas Phase. *J. Electron Spectrosc. Relat. Phenom.* **2012**, *185*, 244–251.
- (41) 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.
- (42) Supplementary materials.
- (43) von Niessen, W.; Schirmer, J.; Cederbaum, L. S. Computational Methods for the One-Particle Green's Function. *Comp. Phys. Rep.* **1984**, *1*, 57–125.
- (44) Cederbaum, L. S.; Domke, W.; Schirmer, J.; von Niessen, W. Correlation Effects in the Ionization of Molecules: Breakdown of the Molecular Orbital Picture. *Adv. Chem. Phys.* **1986**, *65*, 115–159.
- (45) Cederbaum, L. S.; Domke, W. Theoretical Aspects of Ionization Potentials and Photoelectron Spectroscopy: A Green's Function Approach. *Adv. Chem. Phys.* **1977**, *36*, 205–344.
- (46) Ortiz, J. V. Partial Third-Order Quasiparticle Theory: Comparison for Closed-Shell Ionization Energies and an Application to the Borazine Photoelectron Spectrum. *J. Chem. Phys.* **1996**, *104*, 7599–7605.
- (47) Ortiz, J. V. An Efficient, Renormalized Self-Energy for Calculating the Electron Binding Energies of Closed-Shell Molecules and Anions. *Int. J. Quantum Chem.* **2005**, *105*, 803–808.
- (48) Corzo, H. H.; Galano, A.; Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. NR2 and P3+: Accurate, Efficient Electron-Propagator Methods for Calculating Valence, Vertical Ionization Energies of Closed-Shell Molecules. *J. Phys. Chem. A* **2015**, *119*, 8813–8821.
- (49) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.
- (50) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (51) Perdew, J. P. Density-Functional Approximation for The Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B: Condens. Matter* **1986**, *33*, 8822–8824.
- (52) Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. Towards an Order-N DFT Method. *Theor. Chem. Acc.* **1998**, *99*, 391–403.
- (53) Blyth, R. R.; Delaunay, R.; Zitnik, M.; Krempasky, J.; Krempaska, R.; Slezak, J.; Prince, K. C.; Richter, R.; Vondracek, M.; Camilloni, R.; et al. The High Resolution Gas Phase Photoemission Beamline, Elettra. *J. Electron Spectrosc. Relat. Phenom.* **1999**, *101-103*, 959–964.
- (54) Derossi, A.; Lama, F.; Piacentini, M.; Prospero, 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.
- (55) Plekan, O.; Coreno, M.; Feyer, V.; Moise, A.; Richter, R.; de Simone, M.; Sankari, R.; Prince, K. C. Electronic State Resolved PEPICO Spectroscopy of Pyrimidine. *Phys. Scr.* **2008**, *78*, 058105.
- (56) *Baspik 43x63 mm*. [https://baspik.com/eng/products/mkp/dec43\\_63/](https://baspik.com/eng/products/mkp/dec43_63/)
- (57) Torelli, P.; Sacchi, M.; Cautero, G.; Cautero, M.; Krastanov, B.; Lacovig, P.; Pittana, P.; Sergo, R.; Tommasini, R.; Fondacaro, et al. Experimental Setup for High Energy Photoemission Using Synchrotron Radiation. *Rev. Sci. Instrum.* **2005**, *76*, No. 023909.
- (58) Cautero, G.; Sergo, R.; Stebel, L.; Lacovig, P.; Pittana, P.; Predonzani, M.; Carrato, S. A Two-Dimensional Detector for Pump-and-Probe and Time Resolved Experiments. *Nucl. Instrum. Methods Phys. Res. A* **2008**, *595*, 447–459.
- (59) Sa'adeh, H. et al., *Valence PEPICO of Indole and Its Derivatives in the Gas Phase*. (in preparation).
- (60) Kimura, K.; Katsumata, S.; Achiba, Y.; Yamazaki, T.; Iwata, S. *Handbook of HeI Photoelectron Spectra of Fundamental Organic Molecules*; Japan Scientific Societies: Tokyo, 1981.
- (61) Myrseth, V.; Bozek, J. D.; Kukk, E.; Sæthre, L. J.; Thomas, T. D. Adiabatic and Vertical Carbon 1s Ionization Energies in Representative Small Molecules. *J. Electron Spectrosc. Relat. Phenom.* **2002**, *122*, 57–63.
- (62) Petterson, L.; Nordegren, J.; Selander, L.; Nordling, C.; Siegbahn, K.; Ågren, H. Core-Electron Binding Energies in the Soft X-Ray Range Obtained in X-Ray Emission. *J. Electron Spectrosc. Relat. Phenom.* **1982**, *27*, 29–37.
- (63) Thomas, T. D.; Shaw, R. W., Jr. Accurate Core Ionization Potentials and Photoelectron Kinetic Energies for Light Elements. *J. Electron Spectrosc. Relat. Phenom.* **1974**, *5*, 1081–1094.
- (64) Gelius, U.; Allan, C. J.; Johansson, G.; Siegbahn, H.; Allison, D. A.; Siegbahn, K. The ESCA Spectra of Benzene and the Iso-Electronic Series, Thiophene, Pyrrole and Furan. *Phys. Scr.* **1971**, *3*, 237–242.
- (65) Lemmon, E. W.; McLinden, M. O.; Friend, D. G.; Linstrom, P.; Mallard, W. *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; National Institute of Standards and Technology: Gaithersburg, MD, <http://webbook.nist.gov/>
- (66) Calais, J.-L.; Kryachko, E. *Conceptual Perspectives in Quantum Chemistry*; Springer, 1997.
- (67) Ferreira, A. M.; Seabra, G.; Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. Application and Testing of Diagonal, Partial Third-Order Electron Propagator Approximations. In: Cioslowski, J.; Eds.; *Quantum-Mechanical Prediction of Thermochemical Data*; Understanding Chemical Reactivity: Springer, Dordrecht, 2001, pp 131–160.
- (68) Snis, A.; Matar, S. F.; Plashkevych, O.; Ågren, H. Core Ionization Energies of Carbon–Nitrogen Molecules and Solids. *J. Chem. Phys.* **1999**, *111*, 9678–9686.
- (69) Chong, D. P. Computational Study of the Structures and Photoelectron Spectra of 12 Azabenzene. *Can. J. Chem.* **2019**, *97*, 697–703.
- (70) Sankari, R.; Ehara, M.; Nakatsuji, H.; Senba, Y.; Hosokawa, K.; Yoshida, H.; De Fanis, A.; Tamenori, Y.; Aksela, S.; Ueda, K. Vibrationally Resolved O 1s Photoelectron Spectrum of Water. *Chem. Phys. Lett.* **2003**, *380*, 647–653.
- (71) Barber, M.; Bell, D. J.; Morris, M.; Tetler, L. W.; Woods, M. D.; Monaghan, J. J.; Morden, W. E. Mass Spectra of Doubly Charged Ions. *Org. Mass Spectrom.* **1989**, *24*, 504–510.
- (72) Eland, J. H. D.; Feifel, R. *Double Photoionisation Spectra of Molecules*. Oxford University Press, 2018.
- (73) Perreault, H.; Ramaley, L.; Benoit, F. M.; Sim, P. G.; Boyd, R. K. Fragmentation Reactions of Molecular Dications of Aromatic Heterocyclic Fused-Ring Compounds Containing More than One Nitrogen Atom. *Org. Mass Spectrom.* **1992**, *27*, 89–96.
- (74) Carlson, T. A. *Photoelectron and Auger Spectroscopy*; Plenum Press: New York, 1975.
- (75) Püttner, R.; Sekushin, V.; Kaindl, G.; Liu, X.-J.; Fukuzawa, H.; Ueda, K.; Tanaka, T.; Hoshino, M.; Tanaka, H. A Vibrationally Resolved C 1s<sup>-1</sup> Auger Spectrum of CO<sub>2</sub>. *J. Phys. B: At., Mol. Opt. Phys.* **2008**, *41*, 045103.
- (76) Siegbahn, K.; Nordling, C.; Johansson, G.; Hedman, J.; Heden, P. F.; Hamrin, K.; Gelius, U.; Bergmark, T.; Werme, L. O.; Manne, R.; et al. *ESCA Applied to Free Molecules*; North-Holland Publ. Co.: Amsterdam-London, 1971.
- (77) Werner, H.-J.; Knowles, P. J. An Efficient Internally Contracted Multiconfiguration–Reference Configuration Interaction Method. *J. Chem. Phys.* **1988**, *89*, 5803–5814.



(78) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General-Purpose quantum Chemistry Program Package. *Software Focus* **2012**, *2*, 242–253.

(79) Storchi, L.; Tarantelli, F.; Veronesi, S.; Bolognesi, P.; Fainelli, E.; Avaldi, L. The Auger Spectroscopy of Pyrimidine and Halogen-Substituted Pyrimidines. *J. Chem. Phys.* **2008**, *129*, 154309.

(80) Eland, J. H. D. Spectra of the Dications of Benzene, Naphthalene and Azulene. *Chem. Phys.* **2008**, *345*, 82–86.

(81) Molloy, R. D.; Danielsson, A.; Karlsson, L.; Eland, J. H. D. Double Photoionisation Spectra of Small Molecules and a New Empirical Rule for Double Ionisation Energies. *Chem. Phys.* **2007**, *335*, 49–54.

# Supporting Information

## Experimental and Theoretical Photoemission Study of Indole and its Derivatives in the Gas Phase

*Oksana Plekan<sup>1\*</sup>, Hanan Sa'adeh<sup>1,2</sup>, Alessandra Ciavardini<sup>3</sup>, Carlo Callegari<sup>1</sup>, Giuseppe Cautero<sup>1</sup>, Carlo Dri<sup>4</sup>, Michele Di Fraia<sup>1</sup>, Kevin C. Prince<sup>1,5</sup>, Robert Richter<sup>1</sup>, Rudi Sergo<sup>1</sup>, Luigi Stebel<sup>1</sup>, Michele Devetta<sup>6</sup>, Davide Faccialà<sup>6</sup>, Caterina Vozzi<sup>6</sup>, Lorenzo Avaldi<sup>7</sup>, Paola Bolognesi<sup>7</sup>, Mattea Carmen Castrovilli<sup>7</sup>, Daniele Catone<sup>8</sup>, Marcello Coreno<sup>9</sup>, Fabio Zuccaro<sup>9</sup>, Elisa Bernes<sup>10</sup>, Giovanna Fronzoni<sup>10</sup>, Daniele Toffoli<sup>10</sup>, Aurora Ponzi<sup>11</sup>*

<sup>1</sup>Elettra-Sincrotrone Trieste S.C.p.A., in Area Science Park, 34149 Basovizza, Trieste, Italy,

<sup>2</sup>Department of Physics, The University of Jordan, Amman 11942, Jordan,

<sup>3</sup>CERIC-ERIC, in Area Science Park, 34149 Basovizza, Trieste, Italy,

<sup>4</sup>IOM-CNR Laboratorio TASC, Area Science Park, 34149 Basovizza, Trieste, Italy,

<sup>5</sup>Centre for Translational Atomaterials, Swinburne University of Technology, Melbourne 3122, Australia,

<sup>6</sup>Istituto di Fotonica e Nanotecnologie, CNR, Piazza Leonardo da Vinci 32, Milano, Italy,

<sup>7</sup>CNR-Istituto di Struttura della Materia, Area della Ricerca di Roma 1, Monterotondo Scalo, Italy,

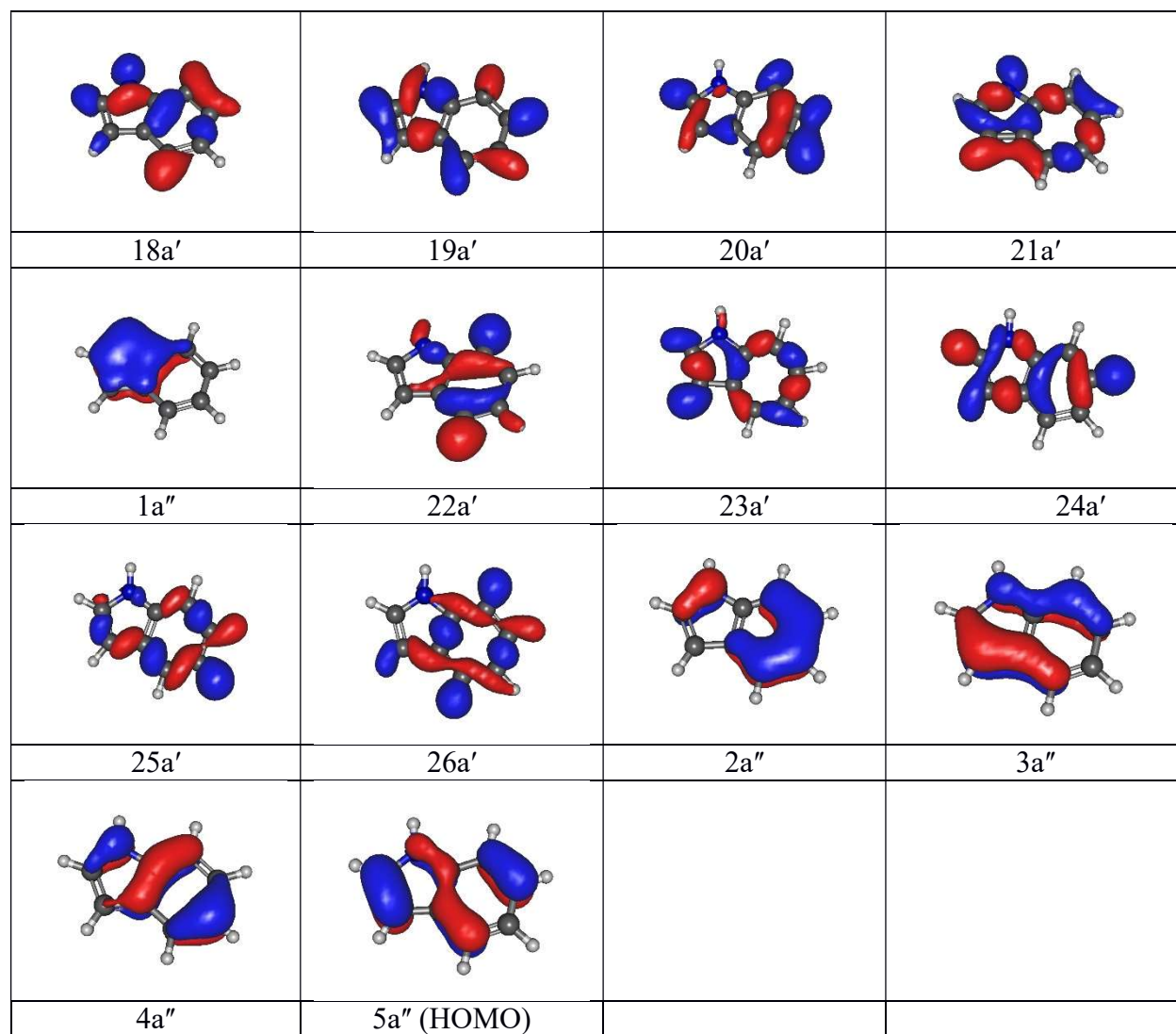
<sup>8</sup>CNR-Istituto di Struttura Della Materia, Area della Ricerca di Tor Vergata, Rome, Italy,

<sup>9</sup>ISM-CNR, in Area Science Park, 34149 Basovizza, Trieste, Italy,

<sup>10</sup>Dipartimento di Scienze Chimiche, Università degli Studi di Trieste, Trieste, Italy,

<sup>11</sup>Institut Ruđer Bošković, Bijenička cesta 54, 10000, Zagreb, Croatia.

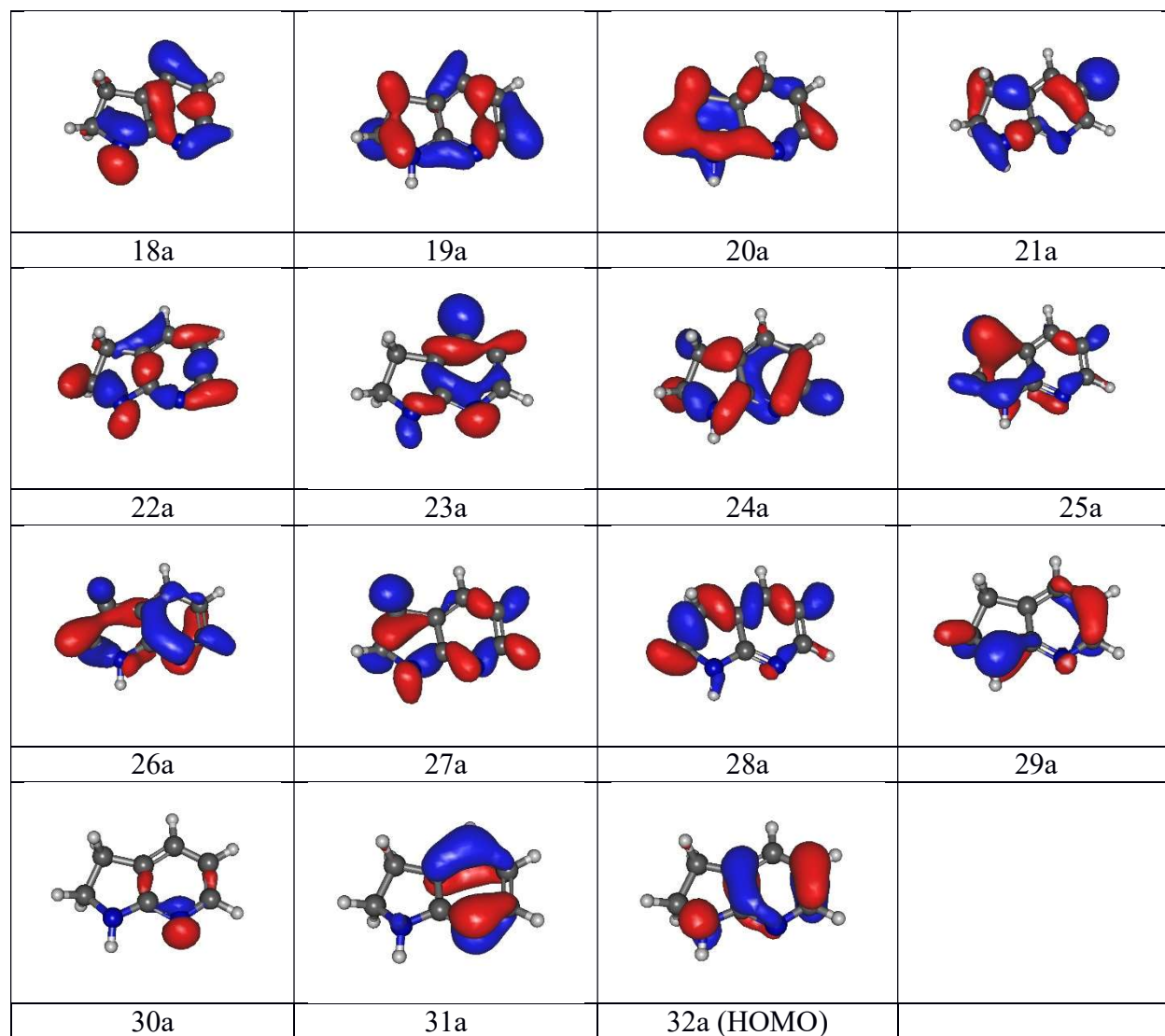
**\*Corresponding author:** E-mail: oksana.plekan@elettra.eu.



**Figure S1.** Plot of HF/cc-pVTZ orbitals for indole (for assignment see Table S1).

**Table S1. Vertical ionization potentials (eV) and pole strengths (PS) of indole calculated with HF, OVGF, P3 and P3+/cc-pVTZ models.**

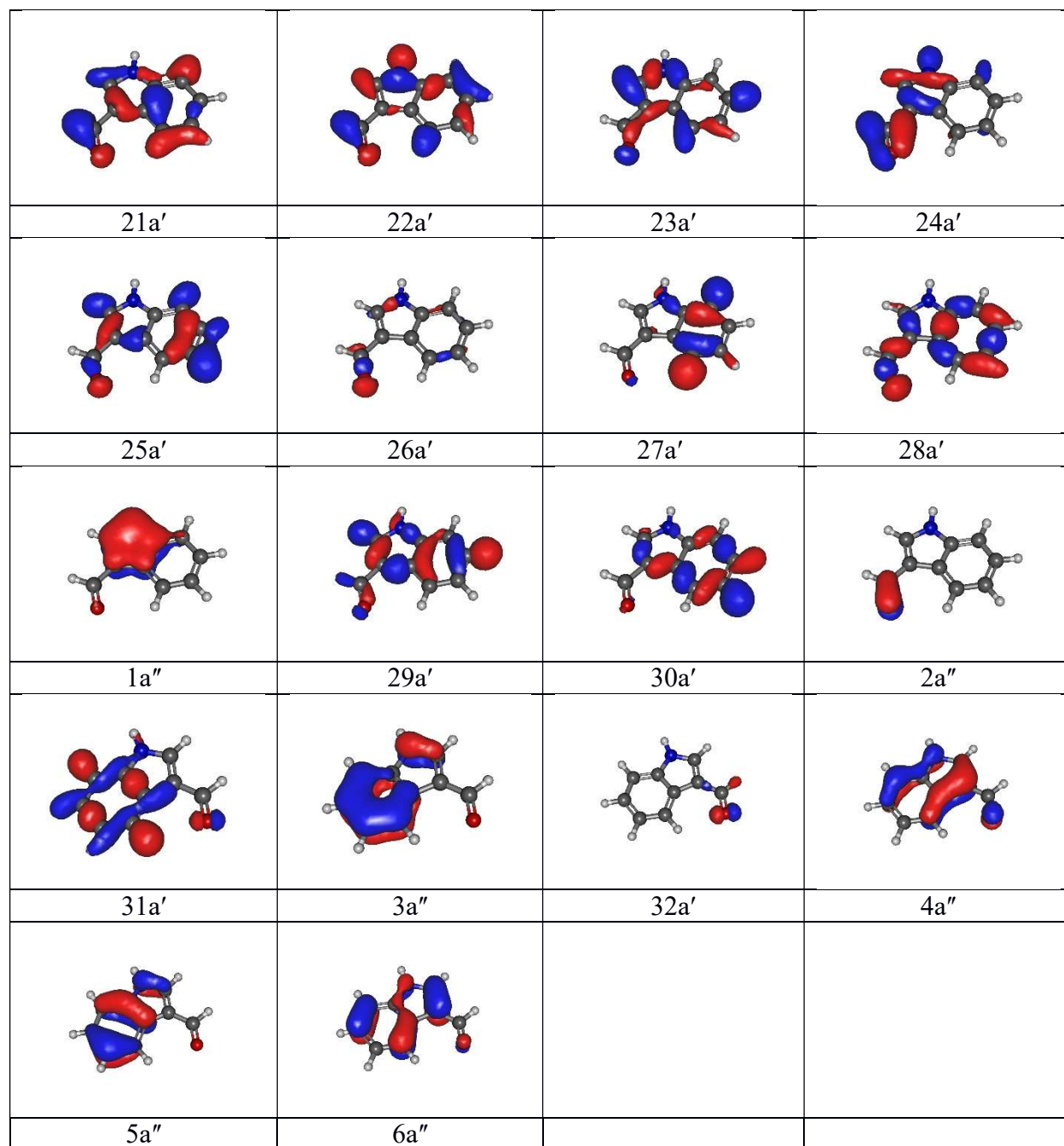
<b>MO</b>	<b>Type</b>	<b>KT</b>	<b>OVGF</b>	<b>PS</b>	<b>P3</b>	<b>PS</b>	<b>P3+</b>	<b>PS</b>
5a''	$\pi$	7.687	7.792	0.894	7.997	0.884	7.911	0.881
4a''	$\pi$	8.158	8.126	0.892	8.341	0.880	8.245	0.877
3a''	$\pi$	10.464	9.811	0.875	10.007	0.863	9.881	0.860
2a''	$\pi$	12.644	11.365	0.815	11.415	0.805	11.285	0.806
26a'	$\sigma$	13.069	11.652	0.896	11.890	0.885	11.677	0.878
25a'	$\sigma$	13.782	12.262	0.890	12.478	0.879	12.270	0.872
24a'	$\sigma$	14.916	13.350	0.884	13.502	0.873	13.291	0.866
23a'	$\sigma$	15.638	13.940	0.881	14.090	0.870	13.878	0.862
1a''	$\pi$	15.852	13.962	0.793	13.866	0.795	13.662	0.793
22a'	$\sigma$	16.177	14.518	0.873	14.571	0.863	14.371	0.856
21a'	$\sigma$	16.349	14.636	0.879	14.786	0.868	14.519	0.858
20a'	$\sigma$	17.616	15.696	0.855	15.762	0.846	15.583	0.840
19a'	$\sigma$	17.961	16.012	0.853	16.102	0.845	15.880	0.838
18a'	$\sigma$	19.795	17.677	0.840	17.717	0.834	17.482	0.827



**Figure S2.** Plot of HF/cc-pVTZ orbitals for 2,3-dihydro-7-azaindole (for assignment see Table S2).

**Table S2. Vertical ionization potentials (eV) and pole strengths (PS) of 2,3-dihydro-7-azaindole calculated with HF, OVGf, P3 and P3+/cc-pVTZ models.**

MO	Type	KS	OVGF	PS	P3	PS	P3+	PS
32a	$\pi$	7.951	7.802	0.897	8.008	0.887	7.916	0.884
31a	$\pi$	10.117	9.614	0.885	9.923	0.873	9.823	0.871
30a	n	11.072	9.496	0.891	9.639	0.879	9.382	0.870
29a	$\pi$	11.751	10.653	0.877	10.678	0.869	10.540	0.865
28a	$\sigma$	13.599	12.325	0.896	12.478	0.887	12.303	0.881
27a	$\sigma$	14.453	13.140	0.888	13.192	0.879	13.035	0.874
26a	$\sigma$	14.612	13.297	0.849	13.285	0.842	13.157	0.841
25a	$\sigma$	14.867	13.620	0.892	13.539	0.884	13.423	0.880
24a	$\sigma$	15.075	13.629	0.860	13.690	0.851	13.517	0.848
23a	$\sigma$	15.429	13.684	0.879	13.789	0.868	13.581	0.861
22a	$\sigma$	16.339	14.688	0.880	14.809	0.869	14.576	0.861
21a	$\sigma$	17.285	15.492	0.867	15.576	0.858	15.388	0.853
20a	$\sigma$	17.792	16.077	0.859	16.083	0.855	15.945	0.853
19a	$\sigma$	18.372	16.459	0.856	16.529	0.849	16.345	0.844
18a	$\sigma$	19.568	17.515	0.855	17.594	0.850	17.365	0.842

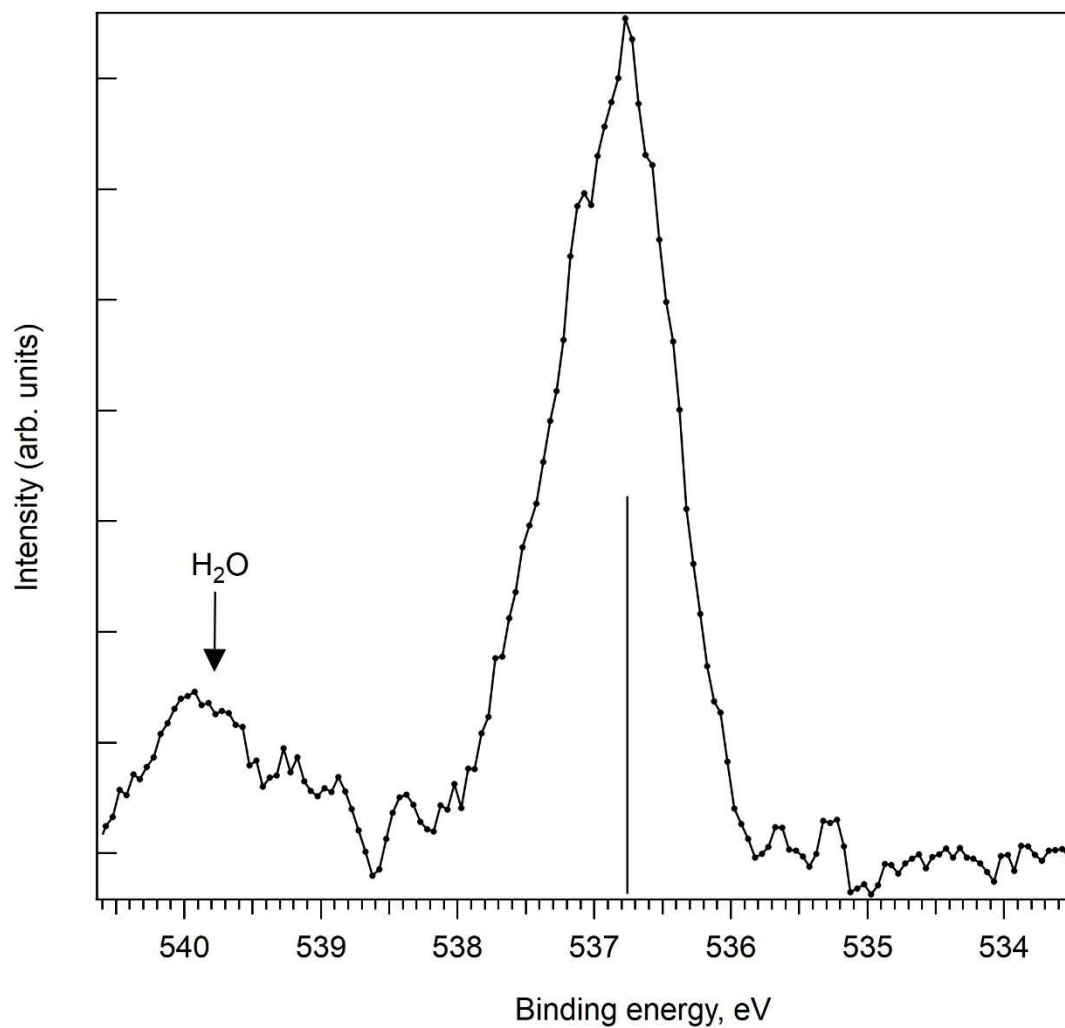


**Figure S3.** Plot of HF/cc-pVTZ orbitals for 3-formylindole with the trans orientation of the formyl group (for assignment see Table S3).

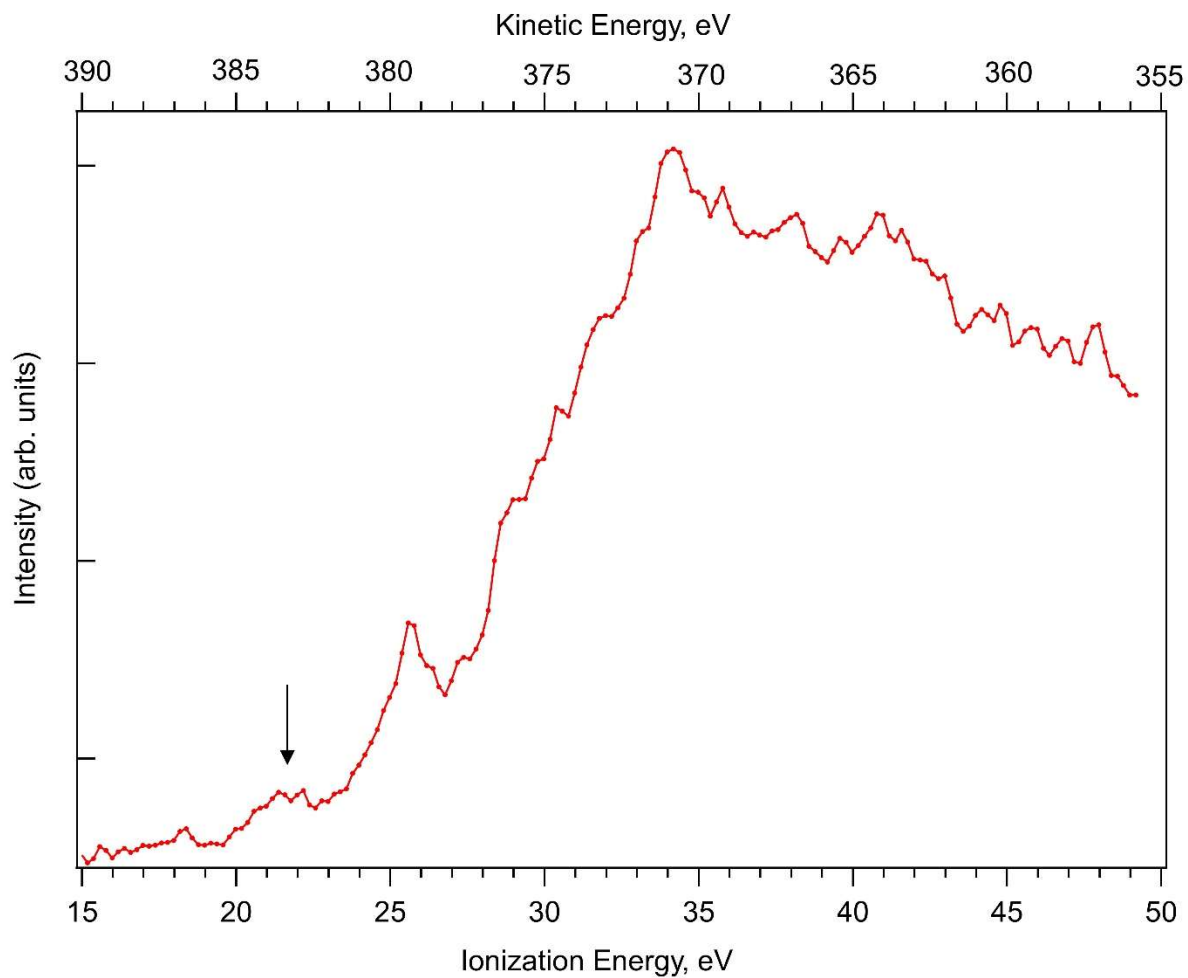
**Table S3. Vertical ionization potentials (eV) and pole strengths (PS) of 3-formylindole calculated with HF, OVGF, P3 and P3+/cc-pVTZ models.**

MO	Type	KT	OVGF	PS	P3	PS	P3+	PS
6a''	$\pi$	8.166	8.178	0.891	8.444	0.881	8.354	0.878
5a''	$\pi$	8.604	8.556	0.888	8.779	0.877	8.684	0.874
4a''	$\pi$	10.678	9.961	0.876	10.234	0.864	10.105	0.860
32a'	n	11.315	9.749	0.871	9.581	0.869	9.339	0.862
3a''	$\pi$	13.029	11.737	0.819	11.794	0.805	11.663	0.805
31a'	$\sigma$	13.571	12.104	0.885	12.254	0.880	12.035	0.873
30a'	$\sigma$	14.023	12.498	0.884	12.694	0.878	12.485	0.871
2a''	$\pi$	14.272	13.226	0.840	13.126	0.829	13.012	0.829
29a'	$\sigma$	15.337	13.620	0.877	13.850	0.871	13.618	0.863
28a'	$\sigma$	16.359	14.398	0.862	14.554	0.860	14.247	0.849
27a'	$\sigma$	16.424	14.766	0.867	14.764	0.862	14.566	0.855
1a''	$\pi$	16.542	14.515	0.793	14.512	0.786	14.308	0.786
26a'	$\sigma$	16.637	14.810	0.865	14.798	0.863	14.514	0.854
25a'	$\sigma$	17.978	15.981	0.854	16.051	0.846	15.851	0.839
24a'	$\sigma$	18.022	15.907	0.853	16.040	0.850	15.775	0.841
23a'	$\sigma$	18.519	16.445	0.850	16.594	0.844	16.342	0.836
22a'	$\sigma$	19.978	17.802	0.848	17.879	0.843	17.637	0.835
21a'	$\sigma$	20.774	18.383	0.834	18.424	0.830	18.184	0.822





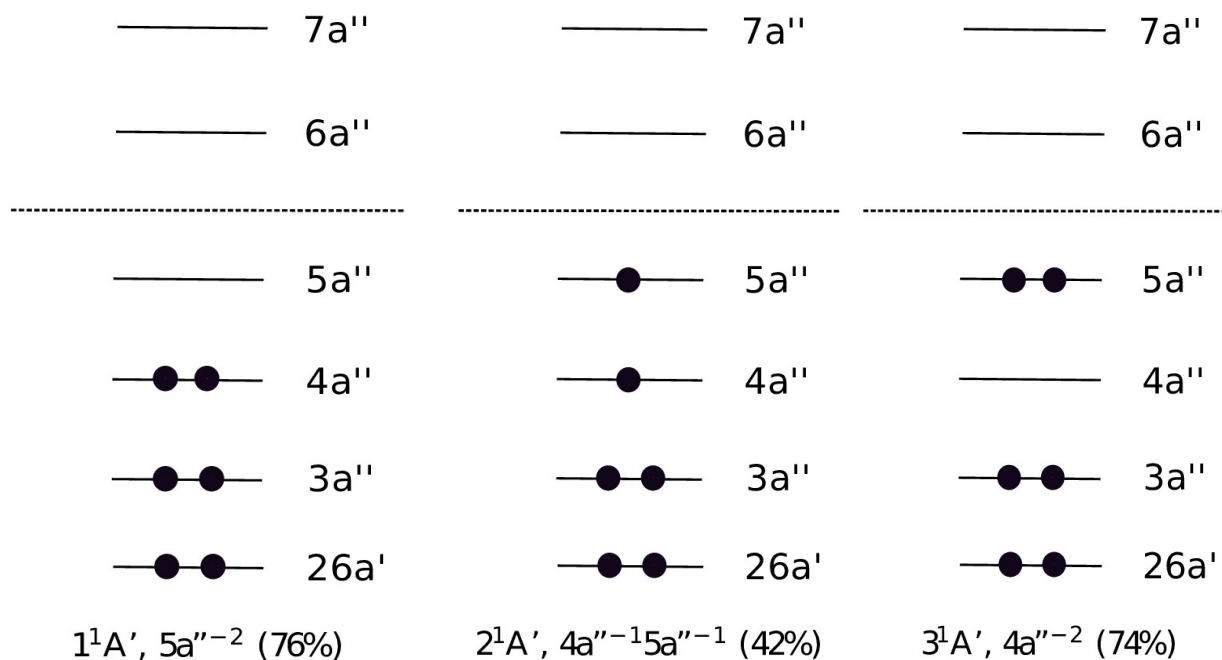
**Figure S4.** O 1s photoemission spectrum of 3-formylindole. Dotted lines: experimental data, bar: theoretical data computed by using the hybrid PW86x potential. Peak at 539.80 eV is due to the water present in the sample and/or the experimental chamber. 1



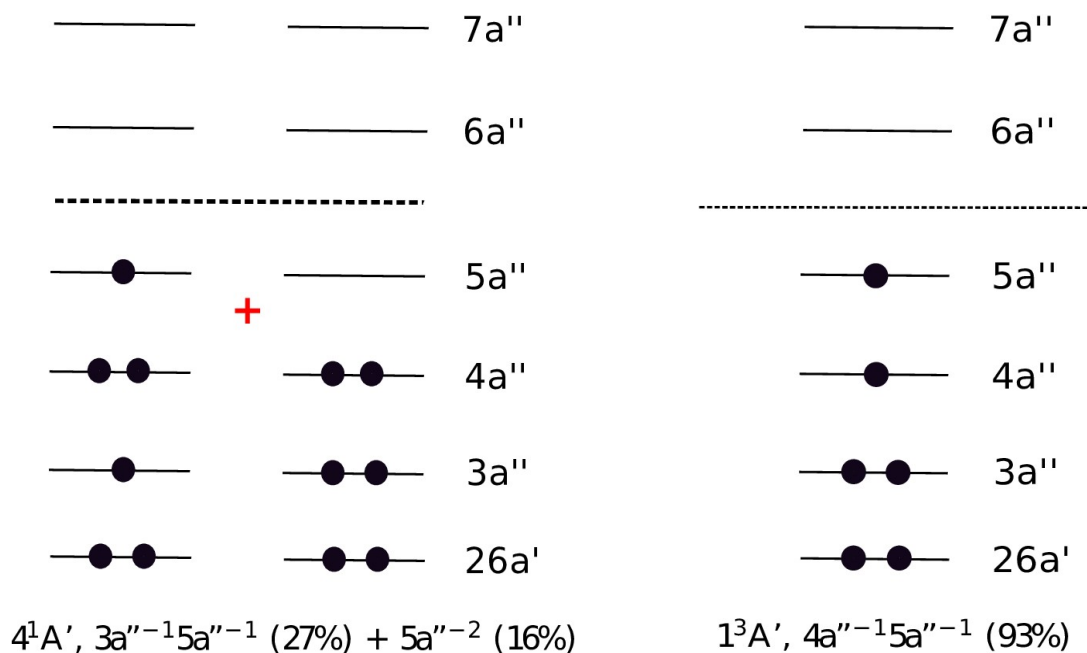
**Figure S5.** N-Auger spectrum of indole measured at photon energy 450 eV.

**Table S4: Calculated double ionization energy values [eV] for indole and the percentage of the dominant reference CI configurations.**

Dication states	MRCI/ANO-VT-TZ	Percentage of the dominant reference CI configurations for the dication states. Note: only percentages greater than 10% have been considered.
$1^1A'$	20.89	$5a''^{-2}$ (76%)
$2^1A'$	21.38	$4a''^{-1} 5a''^{-1}$ (42%)
$3^1A'$	22.42	$4a''^{-2}$ (74%)
$4^1A'$	25.06	$3a''^{-1} 5a''^{-1}$ (27%), $5a''^{-2}$ (16%)
$1^3A'$	20.71	$4a''^{-1} 5a''^{-1}$ (93%)



**Figure S6.** Main CI vectors for double ionized first ( $1^1A'$ ), second ( $2^1A'$ ) and third ( $3^1A'$ ) singlet states of indole. Only configurations greater than 10% have been considered.



**Figure S7.** Main CI vectors for double ionized fourth singlet ( $4^1A'$ ) and triplet state ( $1^3A'$ ) of indole. Only configurations greater than 10% have been considered.

**References:**

- (1) Sankari, R.; Ehara, M.; Nakatsuji, H.; Senba, Y.; Hosokawa, K.; Yoshida, H.; De Fanis, A.; Tamenori, Y.; Aksela, S.; Ueda, K. Vibrationally Resolved O1s Photoelectron Spectrum of Water. *Chemical Physics Lett.* **2003**, *380*, 647–653.

# Electronic Structure Characterization of a Thiophene Benzo-Annulated Series of Common Building Blocks for Donor and Acceptor Compounds Studied by Gas Phase Photoelectron and Photoabsorption Synchrotron Spectroscopies

D. Toffoli,<sup>†</sup> A. Guarnaccio,<sup>\*,‡</sup> C. Grazioli,<sup>‡</sup> T. Zhang,<sup>§</sup> F. Johansson,<sup>§</sup> M. de Simone,<sup>||</sup> M. Coreno,<sup>‡</sup> A. Santagata,<sup>‡</sup> M. D'Auria,<sup>⊥</sup> C. Puglia,<sup>§</sup> E. Bernes,<sup>†</sup> M. Stener,<sup>†</sup> and G. Fronzoni<sup>\*,†</sup>

<sup>†</sup>Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy

<sup>‡</sup>Tito Scalo (PZ) and Trieste LD2 Unit, ISM-CNR, 34127 Trieste, Italy

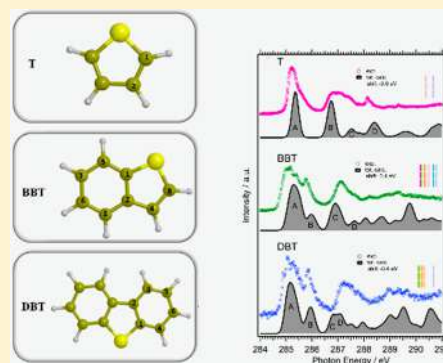
<sup>§</sup>Department of Physics and Astronomy, Uppsala University, Box 516, SE-751 20 Uppsala, Sweden

<sup>||</sup>IOM-CNR, Laboratorio TASC, Sincrotrone Trieste, I-34149 Trieste, Basovizza, Italy

<sup>⊥</sup>Science Department, University of Basilicata, 85100 Potenza, Italy

## Supporting Information

**ABSTRACT:** The near-edge x-ray-absorption fine-structure (NEXAFS) and X-ray photoelectron spectroscopy (XPS) spectra of benzo[*b*]thiophene (BBT) and dibenzothiophene (DBT) in the gas phase have been measured at the carbon K-edge and sulfur L<sub>II,III</sub>-edge regions. The assignment of the spectral features has been provided by theoretical calculations based on density functional theory (DFT) and its time-dependent generalization (TDDFT) in the linear response regime. Observed trends in computed C 1s and S 2p ionization potentials (IPs) have been rationalized in terms of both the inductive effects due to the presence of S and the increased  $\pi$ -electrons delocalization arising from the benzo-annulation process. The analysis of the NEXAFS carbon K-edge and sulfur L<sub>II,III</sub>-edge regions provided information on both low-lying delocalized virtual  $\pi$  orbitals, and higher-lying localized  $\sigma^*(\text{C}-\text{S})$  states. The evolution of the NEXAFS carbon K-edge spectral features along the series thiophene (T) and derivatives, BBT and DBT, is informative of a stabilizing effect due to increased aromaticity. This effect is however more pronounced in going from T to BBT compared to the introduction of a second annulated phenyl ring in DBT. The nature of the most intense sulfur L<sub>II,III</sub>-edge NEXAFS spectral features is instead conserved along the series reflecting thus the localized nature of the virtual states involved in the S 2p core-excitation process.



## 1. INTRODUCTION

Conjugated molecular complexes are of growing interest in the field of organic molecular electronics,<sup>1–5</sup> due to the possibility to combine building block molecules to form more complex structures tuned to the desired performance.<sup>4,5</sup> The introduction of these kinds of molecules in organic solar cells (OSCs) or organic light emitting diodes (OLEDs) has resulted in a further progress of the device efficiency and performance. In a typical OSC, the sunlight energy conversion occurs in photoactive layers/interfaces of donor–acceptor (D–A) materials, often based on two counterparts or, more recently, by combining the electron donor and acceptor properties on two sites of the same molecule usually connected by covalent bonds through  $\pi$ -bridge linkers (D– $\pi$ –A).<sup>6,7</sup> In D– $\pi$ –A systems the choice of the proper  $\pi$ -bridge between the donor D and the acceptor A plays a crucial role not only by adjusting the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) levels but also in the ability of extending the absorption range and improving

the charge separation upon photoexcitation.<sup>8</sup> Hence, characterizing the properties of the isolated building block molecules of donors, acceptors, and  $\pi$ -bridges is essential for predicting, controlling, and manipulating their behavior and applications.

A promising class of electron donors is based on thiophenes<sup>9–11</sup> which are considered as ideal systems being electron rich and providing an outstanding ability to acquire and transport positive charges through a film layer.<sup>9</sup> As a consequence of their efficient light collection, structural versatility and the intrinsic behavior of charge transport, the thiophene-based conjugated systems are considered good candidates for the development of high performance OSCs.<sup>12,13</sup>

Moreover, thiophenes are also considered suitable donors/ $\pi$ -bridges in D– $\pi$ –A molecular systems, capable to improve

Received: August 27, 2018

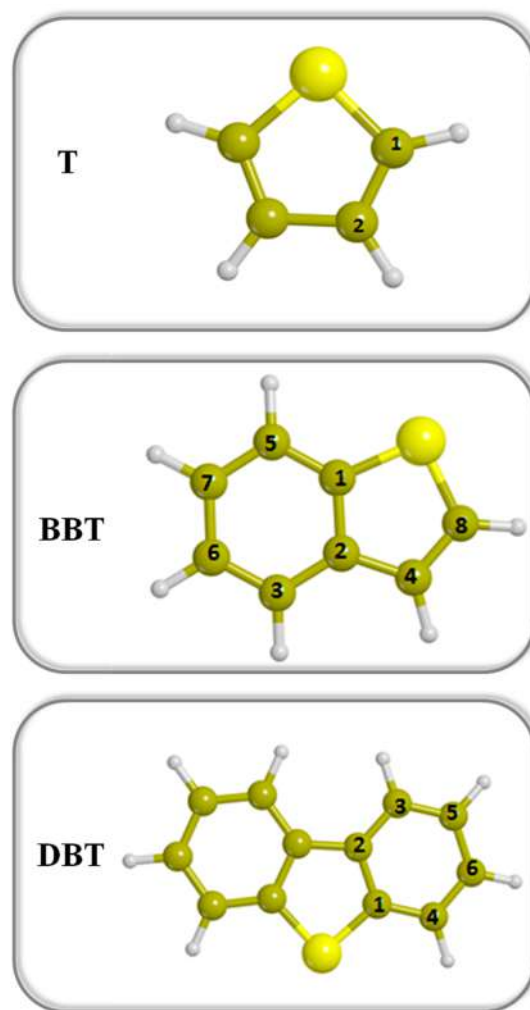
Revised: October 11, 2018

Published: October 15, 2018

the light harvesting efficiency and charge transport properties,<sup>7,14</sup> facilitating the electron migration from the donor to the acceptor counterparts. The introduction of  $\pi$ -conjugated ring-annulated moieties, such as benzene, simultaneously extends the  $\pi$ -conjugation and improves the stability of the resulting molecule. In particular, benzo-annulation at the 2,3-position of thiophene rings results in a class of promising building blocks to be used as electron-donors and  $\pi$ -bridges for the assembling of organic electronic materials.

The present work focused on the electronic structure characterization of building block molecules such as benzo[*b*]thiophene (BBT) and dibenzothiophene (DBT) representing planar and rigid thiophene benzo-annulated structures with extended  $\pi$ -conjugated systems. These structural elements allow greater intermolecular overlap in thin films and consequently provide high charge carrier mobility.<sup>15</sup> Indeed, previous studies show that the systems containing annulated-ring planar structures improve the device performance.<sup>16</sup> As an example, it can be mentioned the 2,8-bis(diphenylphosphoryl)-dibenzo[*b,d*]thiophene (PPT) that is formed by two phosphine oxide moieties functionalizing the small dibenzothiophene core. This derivative is characterized by high triplet energy and is known as good vacuum sublimable electron transporting host material for blue OLEDs. Considering the structure of the PPT molecule, it has been previously demonstrated that the phosphine oxide groups act as breaking points of  $\pi$ -conjugation between the core and the outer groups leaving the electronic structures of the compound practically matching those of the central dibenzothiophene moiety.<sup>17</sup> This means that the phosphine oxide moiety has practically no impact on the core dibenzothiophene geometry and its lowest triplet energy but only on the ionization potential and electron affinity of the resulting system. The performance of devices in which such compounds are used, is especially improved when, among other analogues cores (i.e., carbazole and dibenzofuran), the dibenzothiophene is used as  $\pi$  conjugated core.<sup>18</sup> For this reason, we focused our study on the DBT and its simpler analogue BBT, to understand how the increasing benzo-annulation of the core system influences the electronic property of bigger compounds, like PPT.

Until now, there is only a limited amount of studies about benzo- and dibenzothiophene-based systems whose electronic properties have been described by combined experimental and DFT core–electron excitations.<sup>19–23</sup> To better assess the role and the electronic properties of the thiophene-based building blocks, in this study we investigate both the dibenzothiophene (DBT) and the smaller benzo[*b*]thiophene (BBT) moieties and compare their behaviors with the well-known features of thiophene (T)<sup>24–27</sup> (see Figure 1 for all the chemical structures). The scope is to characterize the electronic structure of the electron-rich T alone and the modification induced by the benzo-annulated rings such as in BBT and DBT. For this purpose, the electronic structures of BBT and DBT in the gas phase have been probed by means of X-ray photoelectron spectroscopy (XPS) and near-edge x-ray absorption fine structure (NEXAFS) spectroscopy. Core–electron spectroscopies are useful tools to investigate the electronic structure of free molecules due to the localized character of the core hole.<sup>28</sup> NEXAFS is solidly accepted as site-sensitive probe of the spatial distribution and local symmetry of the empty states<sup>29</sup> and the closely related XPS, which measures the ionization energies (IEs) of the core electrons, is very sensitive to the local chemical and physical



**Figure 1.** Chemical structures of thiophene (T), benzo[*b*]thiophene (BBT), dibenzothiophene (DBT) molecules.

environment of the ionized atomic site. The observed C 1s and S 2p binding energy (BE) variations have been explained in terms of both the inductive effects due to the presence of S and the increased  $\pi$ -electrons delocalization arising from the annulation process.

The NEXAFS experimental spectra have been collected at both the C K-edge and S L<sub>II,III</sub>-edge and have been rationalized with theoretical calculations performed in the framework of the density functional theory (DFT) approach, employing different and well established computational protocols depending on the not degenerate (K-shell) /degenerate (L<sub>II,III</sub> shells) nature of the core hole. These combined NEXAFS studies have enabled us to completely characterize the low-lying virtual orbitals of the molecules in terms of the atomic C p and S s, d components.

The acquired knowledge about the isolated molecules is useful to predict, control and manipulate their behavior when introduced in bigger molecular systems as for example in PPT or in D– $\pi$ –A complexes.

## 2. EXPERIMENTAL SECTION

Commercially available Sigma-Aldrich benzo[*b*]thiophene, BBT, (purity 98%, mp 30–33 °C) and dibenzothiophene, DBT, (purity 98%, mp 97–100 °C) were used for the

experimental characterizations. BBT and DBT are crystalline solids and were sublimated in vacuum using a custom built resistively heated furnace. DBT was sublimated at 44 °C after purification at around 30 °C for 12 h. BBT is a solid compound with low melting point, and its best vaporization temperature was found to be 28 °C after a quick purification at this temperature for a few hours.

The measurements on the gas phase samples were performed at the gas phase beamline of the Elettra synchrotron in Trieste<sup>30</sup> using a Scienta SES-200 electron analyzer<sup>31</sup> mounted at the magic angle with respect to the electric field vector of the linearly polarized incident light.

C 1s XP spectra of BBT and DBT molecules were recorded at photon energies of 382 and 392 eV, with an overall energy resolution of about 200 and 100 meV, respectively. C 1s XP spectra have been calibrated with respect to the C 1s binding energy (BE) of CO<sub>2</sub> (297.6 eV).<sup>32</sup>

S 2p XP spectra of BBT and DBT molecules were recorded at a photon energy of 260 eV with an overall energy resolution of about 180 and 94 meV, respectively and they were calibrated with respect to the S 2p binding energy (BE) of SF<sub>6</sub> (180.21 eV for 2p<sub>3/2</sub> and 181.5 eV for 2p<sub>1/2</sub>).<sup>33,34</sup>

The C 1s and S 2p XP peaks were fitted using IGOR PRO; we used an in-house written curve fitting routine to reproduce the line-shape of C 1s, while for S 2p we used the package XPST (X-ray Photoelectron Spectroscopy Tools) by Dr. Martin Schmid.<sup>35</sup>

NEXAFS spectra at the C K-edge and S L<sub>II,III</sub>-edge were acquired by measuring the total ion yield (TIY) with an electron multiplier placed in the experimental chamber in front of the ionization region. The photon flux was measured simultaneously using a calibrated Si photodiode (AxVU100 IRD) for the spectra normalization. The energy scale of the C K-edge spectra was calibrated by taking simultaneous spectra of the samples and of CO<sub>2</sub>, with the characteristic transition at 290.77 eV (C 1s → π\*, CO<sub>2</sub>).<sup>36,37</sup> The photon energy resolution was around 110 meV for BBT and 65 meV for DBT molecules.

The energy scale of the S L<sub>II,III</sub>-edge NEXAFS spectra was calibrated by taking simultaneous spectra of the samples and of SF<sub>6</sub>, with the characteristic transition T<sub>1u</sub> (a<sub>1g</sub>)-3/2 T<sub>1u</sub> (a<sub>1g</sub>)-1/2 (at 172.5 and 173.6 eV, respectively).<sup>38</sup> The photon energy resolution was about 80 meV for BBT and about 37 meV for DBT spectra.

### 3. COMPUTATIONAL DETAILS

The equilibrium geometries of thiophene (T), benzo[*b*]-thiophene (BBT), and dibenzothiophene (DBT) have been optimized at the density functional theory (DFT)<sup>39</sup> level, within the local-density approximation (LDA) by using the Vosko, Wilk, and Nusair (VWN) exchange and correlation (xc) functional<sup>40</sup> and the triple ζ polarized (TZP) basis set of Slater type orbitals (STOs) taken from the ADF (Amsterdam density functional) database.<sup>41,42</sup> Optimized geometries obtained with the VWN xc functional in combination with the TZP basis set are in rather good agreement with the experimental geometries of T and DBT derived from X-ray diffraction data.<sup>43,44</sup> (the comparison with the experimental data is reported in Tables S1 and S2 of the [Supporting Information](#)). The three molecules are planar (C<sub>2v</sub> point-group symmetry for T and DBT, C<sub>s</sub> point-group for BBT), and the full symmetry of the systems has been exploited in the calculations.

The C K-edge NEXAFS spectrum for each nonequivalent C site has been computed at the DFT level within the generalized gradient approximation (GGA) for the XC energy functional with the PW86xPerdew functional<sup>45</sup> and the transition potential (TP) scheme.<sup>46,47</sup> In this approach, the Kohn–Sham (KS) orbitals are determined self-consistently with an orbital occupation scheme obtained by removing half an electron from the excited core orbital. Relaxation effects upon formation of the core hole are usually adequately described.<sup>48</sup> The set of virtual molecular orbitals (MOs) obtained from a single TP calculation is used to treat the entire manifold of excited states. To accurately describe excitation to diffuse Rydberg states close to the ionization threshold, an even tempered quadruple-ζ with three polarization and three diffuse functions (designed as ET-QZ3P-3DIFFUSE set in the ADF database) basis set has been used for the core-excited C atom, while a TZP basis has been employed for the remaining atoms. In particular, a frozen core TZP.1s basis set has been employed for the C atoms, and a TZP.2p basis for S. Beside ensuring the localization of the half core hole, this basis set allows for a good accuracy with a further reduction of the computational cost.

Within the TP scheme, excitation energies are obtained as eigenvalues differences between the virtual orbital and the 1s core orbital

$$\Delta E_{i \rightarrow f} = \epsilon_f^{TP} - \epsilon_i^{TP} \quad (1)$$

while transition intensities are expressed in terms of oscillator strengths,  $f_{i \rightarrow f}$ . For samples in the gas phase

$$f_{i \rightarrow f} = \frac{2}{3} n_i \Delta E_{i \rightarrow f} |\langle \varphi_f^{TP} | \mu | \varphi_i^{TP} \rangle|^2 \quad (2)$$

involving dipole matrix elements between initial and final TP MOs while  $n_i$  is the occupation number of the core orbital in the ground state. The Ionization Potential (IP) is defined as the negative of the TP eigenvalue related to the initial core orbital,  $IP = -\epsilon_i^{TP}$ , but, since absolute transition energies are usually too large (the TP approach leads generally to a less attractive potential), a better estimate is obtained by first computing the IPs at the ΔKS (ΔSCF Kohn–Sham) scheme, allowing a full relaxation of the ionized core hole, and shifting the TP excitation energies eq 1 by an amount given by the energy difference  $\epsilon_i^{TP} - IP^{\Delta KS}$ .

In the C K-edge NEXAFS spectra calculations, a separate computation of the excitation spectrum of each nonequivalent C site has been performed, and the total spectrum has been obtained by summing up the different contributions. For each spectrum, theoretical C 1s ionization thresholds are used to separate the below- and the above-edge regions. Only transitions which occur below the ionization threshold can be accurately described by the employed computational protocol, while above it only qualitative information can be extracted, since the electronic continuum wave function cannot be properly described with standard basis sets of quantum chemistry programs.

The spectral structures result from the electronic excitations from a core orbital, deeply localized on the absorbing atom, toward virtual molecular orbitals (MOs), lying around the ionization limit of that specific core hole. The excitation process is dictated by dipole selection rules, and the oscillator strengths are directly connected with the atomic site component of the virtual orbitals which is dipole allowed. Therefore, the K-edge structures are dominated by the  $s \rightarrow p$

dipole-allowed transitions; hence, the transition intensities map the C p components in the final MOs. The features observed at the S  $L_{II,III}$  edges involve the  $2p \rightarrow s/d$  electric-dipole transitions which therefore characterize the final MOs with significant S s, d components.

The  $\Delta$ KS C 1s IPs are also used to simulate the XP spectra. Preliminary calculations of  $\Delta$ KS IP values for both BBT and DBT by using the hybrid B3LYP xc potential<sup>49,50</sup> were performed and the results were in good agreement with those computed with the PW86xPerdew xc. The comparison of the results is reported in Table S3 of the Supporting Information. The accurate calculation of the XPS intensity would in principle require the calculation of the electronic continuum wave function at the given photoelectron kinetic energy. Since the XP spectra are measured with a photon energy well above the C 1s ionization threshold, at good approximation the XPS intensity for each nonequivalent site  $C_i$  is taken to be proportional to the number of equivalent centers in the molecule (2 for T and DBT, 1 for BBT). Line spectra are then convoluted with Gaussian functions with full-width at half-maximum (fwhm) of 0.30 eV. Moreover, a rigid shift of 0.10 eV (for BBT and DBT) and 0.20 eV (for T) to lower BEs has been applied to the theoretical profiles for a better comparison with the experimental data.

The computational protocol used for the calculation of the S  $L_{II,III}$ -edge NEXAFS spectra is more elaborate than that used for the calculation of the C K-edge due to the need of an explicit inclusion of both the coupling between different excitation channels from the S 2p degenerate core-holes and spin-orbit coupling effects.<sup>27,51,52</sup> Therefore, the S  $L_{II,III}$ -edge NEXAFS spectra calculations have been performed within the Time Dependent DFT (TDDFT) by using the two-component zeroth-order regular approximation<sup>53</sup> (ZORA), implemented in the ADF code<sup>41,42,54</sup> within the adiabatic local density approximation (ALDA) to the XC kernel. Here we will only give an outline of the TDDFT formalism for core electron excitations, referring to previous works<sup>52,55</sup> for a more detailed account.

In linear response TDDFT,<sup>56</sup> transition energies and intensities are obtained through the solution of the following eigenvalue equation:

$$\Omega F_i = \omega_i^2 F_i \quad (3)$$

where the elements of the  $\Omega$  matrix are given by

$$\Omega_{ia\sigma,bj\tau} = \delta_{\sigma\tau} \delta_{ij} \delta_{ab} (\epsilon_a - \epsilon_i)^2 + 2\sqrt{(\epsilon_a - \epsilon_i)} \frac{\partial F_{ia}}{\partial P_{jb}} \sqrt{(\epsilon_b - \epsilon_j)} \quad (4)$$

In eq 4, indices  $i$  and  $j$  run over the set of occupied spinors in the KS ground-state, while indices  $a$  and  $b$  run over the set of virtual spinors;  $\epsilon_i$  and  $\epsilon_a$  are the KS molecular orbital energies.  $F$  and  $P$  represent the Fock matrix and the density matrix, respectively, whereas  $\frac{\partial F_{ia}}{\partial P_{jb}}$  are the elements of the coupling matrix. Eigenvalues  $\omega_i^2$  in eq 1 correspond to the square of the excitation energies, while the oscillator strengths can be extracted from the eigenvectors  $F_i$ .

Since core excitation energies lie very high in the excitation spectrum, it is not feasible to extract the corresponding eigenvalues and eigenvectors of the  $\Omega$  matrix by using Davidson's iterative algorithm which is instead very efficient to extract the lowest states.<sup>57</sup> To overcome this issue, the

core-valence separation approximation (CVS) introduced by Cederbaum et al.<sup>58,59</sup> is invoked whereby the configuration space spanned by the solutions of eq 3 (1h-1p space of the single excitations from the KS reference determinant) can be reduced by keeping only configurations that are mandatory for an accurate description of the phenomenon, in the same spirit of *ab initio* configuration interaction (CI) calculations.<sup>60</sup> Therefore, in the application of TDDFT to core excitations, the indexes spanning the occupied orbitals space ( $i$  and  $j$ ) are limited to run only over the core-shell under study, and the eigenvalue eq 3 is solved efficiently with Davidson's algorithm for the lowest roots of this submatrix. In the present case of 2p excitations, the indices of occupied spinors run over the two subshells  $2p_{1/2}$  and  $2p_{3/2}$ , hence allowing a mixing of only those initial states.

The all-electron TZP basis set, optimized for ZORA calculations,<sup>61</sup> and taken from the ADF database has been employed for C and H atoms. For the S atom, the QZ4P ZORA basis set has been enlarged by adding two shells of s, p, d, and f diffuse functions with exponents obtained with the even-tempered criterion ( $\beta = 1.7$ ). The final basis set is therefore the [13s, 9p, 5d, 6f] set, and denoted ET-QZ4P-2diff. The use of a basis set enriched with the inclusion of diffuse functions on the excited atom allows for an improved representation of the relaxation effects of the inner orbitals as well as for a proper description of the higher energy excitations which contribute to the near-edge structures.<sup>62</sup> For a meaningful comparison with the experiment, NEXAFS spectra are plotted after a convolution of the calculated discrete lines with Gaussian functions of fwhm of 0.30 eV.

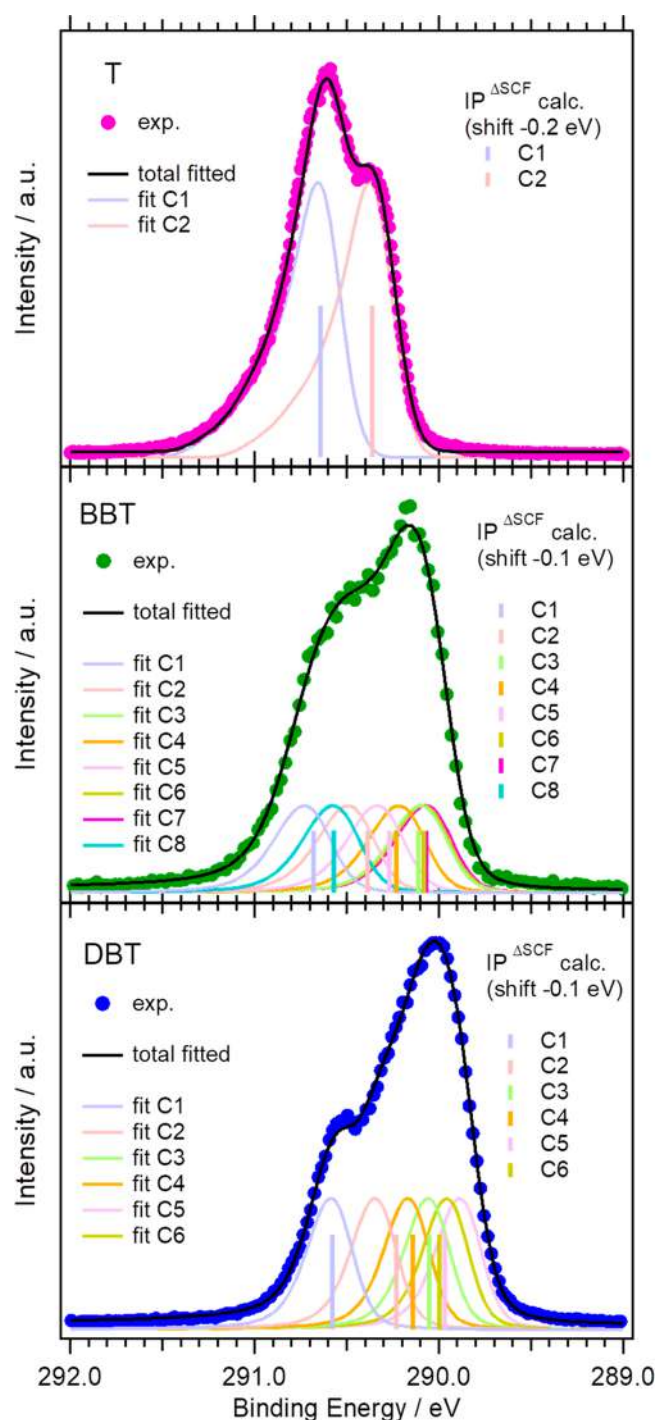
In the self-consistent field (SCF) calculations, the LB94 xc potential by van Leeuwen and Baerends<sup>63</sup> has been employed with the ground state electron configuration; this choice is justified by the correct asymptotical behavior of this xc potential, which is fundamental to properly characterize the high energy valence virtual orbitals and Rydberg states. Scaled ZORA spinor eigenvalues have been used in the calculations of core excitation energies, while IPs are taken as the negative of the LB94 ZORA eigenvalues,  $IP = -\epsilon_{2p}$ , where  $\epsilon_{2p}$  represents the eigenvalue of the 2p orbital. Calculated IPs have been shifted to lower BEs by 0.28 (T), 1.85 (BBT), and 0.90 eV (DBT) for a better comparison with the experimental measurements.

## 4. RESULTS AND DISCUSSION

In the following subsections we will discuss the TP-DFT and TDDFT results to obtain an assignment of the experimental XP and NEXAFS spectra of BBT and DBT. In the first two subsections, we will consider the C 1s and S 2p XP spectra, while the corresponding NEXAFS spectra will be the subject of the final two subsections. The spectra of thiophene (T) from ref<sup>26,27</sup> measured under similar experimental conditions will be also reported for comparison.

**4.1. C 1s XPS.** Photoelectron spectra of the C 1s core levels are displayed in Figure 2 together with the analysis of the experimental line-shapes performed with the fitting procedure described in the Experimental Section. The following analysis is based on the fact that the core level BEs are very sensitive to the chemical environment of the ionized atomic site, in terms of electron density (initial state effect) and electronic relaxation (final state effect) which both contribute to the BE chemical shifts. From a general overview, the C 1s XPS line profiles of BBT and DBT appear quite similar, presenting a





**Figure 2.** C 1s XPS experimental data (dots lines) of T (top panel), BBT (middle panel), and DBT (bottom panel) are shown together with the total fitted line-shape (black solid profiles) and the  $C_i$  peaks (solid colored lines) resulting from the fitting procedure using Voigt curves with 0.26 eV fwhm (for BBT) and 0.30 eV fwhm (for DBT). Moreover, a comparison with the theoretical IPs (solid colored bars) calculated at  $\Delta$ KS level is also presented; their intensities are proportional to the multiplicity of  $C_i$  atoms. The calculated lines reported in the graphs have been shifted by  $-0.1$  (for BBT and DBT) and  $-0.2$  eV (for T) in order to match the experimental results.

peak of higher intensity at about 290.16 eV (BBT) and 290.02 eV (DBT), and a lower intensity shoulder at 290.53 and 290.55 eV, respectively. However, as expected, the two spectra show also different broadening, intensity variations and peak

shifts which, as discussed later, are related to the molecular structures.

A useful comparison of the XP spectra of the BBT and DBT molecules together with the smaller T precursor is reported in Figure 2 (top, middle and bottom panels, respectively for T, BBT and DBT), enlightening mainly two effects as responsible for the spectral line-shapes: the symmetry variation and the increase of the molecular complexity along the series T-BBT-DBT, with different number of C atoms, both contributing to the peak broadening.

Since the C 1s XP spectrum of T has been already discussed in details in a previous work,<sup>26</sup> in the following discussion we will only focus on BBT and DBT.

It is known that the C 1s photoelectron spectra of thiophenes are affected by a pronounced vibrational envelope<sup>24</sup> that has a strong effect on the spectrum of T and then becomes weaker for the more complex molecules.<sup>26</sup> The vibrational tail, however, is hardly discernible in the present BBT and DBT measurements. The fitting of the experimental data, considering the different kinds of C atoms ( $C_i$ ) of the molecules, have allowed the identification of the chemical shifted contributions (BEs) to be compared with the theoretical IPs. The results are shown in Figure 2 and the experimental fitted BEs (BEs fit) and the theoretical IPs (IPs calculated and shifted) are reported in Table 1.

**Table 1.** Comparison between Theoretical  $\Delta$ KS C 1s IPs and Experimental BEs, Derived from the Fitting Described in the Text for All Non-Equivalent  $C_i$  Atoms in BBT and DBT<sup>b</sup>

peaks	BBT		DBT	
	$\Delta$ KS IPs <sup>a</sup>	BEs (fit)	$\Delta$ KS IPs <sup>a</sup>	BEs (fit)
C1	290.68	290.73	290.58	290.59
C8	290.57	290.58	—	—
C2	290.39	290.49	290.23	290.35
C5	290.27	290.33	289.97	289.89
C4	290.23	290.22	290.14	290.17
C3	290.11	290.10	290.05	290.06
C6	290.08	290.09	290.00	289.96
C7	290.07	290.07	—	—

<sup>a</sup>Calculated ionization potentials shifted by  $-0.10$  eV to match the experimental peaks. <sup>b</sup>All values are expressed in eV.

For BBT (middle panel of Figure 2), the calculations assign the highest C 1s IP to the C1 atom shared by the two fused rings of the molecule. This first peak is separated by 110 meV from the IP of C8, which is also bonded to sulfur as well as to C4 and to a hydrogen atom. This theoretical predicted chemical shift between C1 and C8 is in good agreement with the experimental value of 150 meV. Progressively lower IPs are assigned to C2, C5, and C4 atomic sites, while the IPs of the remaining C atoms of the phenyl ring (C3, C6, and C7) are all clustered in a narrow energy range.

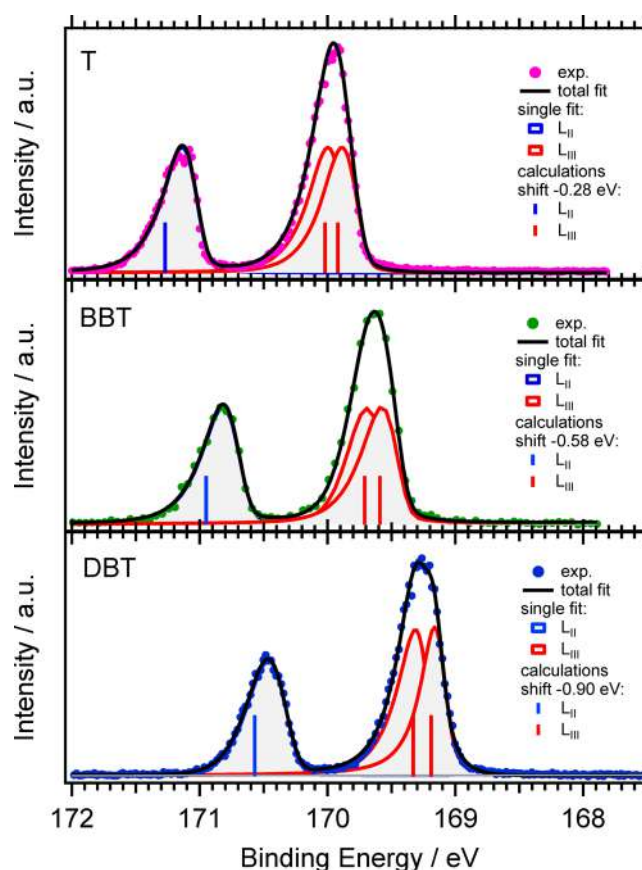
The observed trend can be rationalized by considering two effects: the higher electronegativity of S compared to C, as well as the charge transfer from S to C atoms not bonded with it through resonance in the  $\pi$  system. Moreover, the higher IP of C1 compared to C8 and the progressively lower IP of the external C atoms (C3–C7) resemble a trend already observed in polycyclic aromatic hydrocarbons for which higher IPs are expected for C atoms not bonded to any H atom compared to external carbons bound to H atoms.<sup>64</sup> A comparison of the

$\Delta$ KS C 1s IPs with the experimental BEs, reported in Table 1, reveals a general good agreement between theory and experiment, since the discrepancies between the two sets of data are of the order of 0.1 eV or less, giving us confidence in the proposed assignment.

A similar good agreement between theory and experiment is found for DBT. Also in this case, the highest C 1s IP is associated with the C1 atom well separated, by 350 meV, from the IP of C2. The experimental chemical shift (estimated by the fit) between the C 1s XP lines of C1 and C2 is 240 meV. We note that the theory slightly overestimates this chemical shift, by 110 meV; however, this difference is well within the errors of the computational protocol used. The IPs of the remaining C atoms follow a trend similar to what was already discussed for BBT, whereby lower IPs are associated with external C atoms of the phenyl ring (C3–C6) bonded to H atoms. The BE energy sequence of the Ci of DBT is consistent with an inductive effect of the electronegative S atom and of the electron delocalization through the molecular  $\pi$  system.

Considering T alone (already discussed deeply in ref 26) it can be noticed that the agreement among calculated IPs (290.86 and 290.58 eV,  $\Delta E_{IP} = 288$  meV) and fitted BEs (290.60 and 290.35 eV,  $\Delta E_{BE} = 250$  meV), respectively for C1 and C2 was quite good. Even better is the accord among the chemical shifts estimated by data fit compared to their corresponding calculated IPs values for BBT and DBT as reported in Table 1. In details, the chemical shift between the BEs of the C1 and C2 sites remains constant for both BBT and DBT ( $\Delta E_{BE} = 240$  meV), while the calculated chemical shift of the IPs of C1 and C2 are  $\Delta E_{IP(BBT)} = 290$  meV for BBT and  $\Delta E_{IP(DBT)} = 350$  meV for DBT. Moreover, we noticed that the experimentally estimated (by fit) chemical shifts between C4 and C8 of BBT are  $\Delta E_{BE} = 360$  meV and the theoretical value is  $\Delta E_{IP} = 340$  meV. From BBT to DBT, we introduce one more annulated phenyl ring, which can contribute to stabilize more the C2 and C1 sites in DBT than in BBT by giving a more symmetric molecular geometry and a more extended  $\pi$  system. Furthermore, the addition of a second annulated phenyl ring also causes a decrease of the BEs of the external C atoms of the phenyl rings (in the range of 0.04–0.2 eV), confirming an increase of charge density for aromatic conjugation of the system going from BBT to DBT. Indeed, from a chemical point of view, all the molecules under investigation are aromatic according to Hückel rule ( $4n + 2$ ). In particular, along the series T-BBT-DBT there is an increase of charge density due to the  $\pi$  delocalized electrons: for T (6  $\pi$  electrons)  $n = 1$ , for BBT (10  $\pi$  electrons)  $n = 2$ , and for DBT (14  $\pi$  electrons)  $n = 3$ .

**4.2. S 2p XPS.** Photoelectron spectra of the S 2p core levels of BBT, DBT and, for comparison of T, are reported in Figure 3, together with the analysis of the experimental line-shapes performed by the fitting procedure (described in the Experimental Section). The related experimental fitted BEs (obtained by a single fit of each component) and the theoretical IPs (calculated and shifted) of each S 2p components are reported in Table 2. Also the data for T from ref 27 are reported in Table 2 to facilitate the comparison. The S 2p core hole splits into two components, namely  $2p_{3/2}$  ( $L_{III}$ ) and  $2p_{1/2}$  ( $L_{II}$ ), due to the spin–orbit (SO) interaction. Moreover, the anisotropic molecular field determines a further splitting of the degenerate  $2p_{3/2}$  level into two components which transform according to the irreducible representations of the point group of the molecule.



**Figure 3.** S 2p XP spectra of T (top panel), BBT (central panel), and DBT (bottom panel): experimental data (dots) are shown together with the results of the total line-shape of fitting procedure (black line). The light gray filled curves are the fitted  $L_{II}$  (blue lines) and  $L_{III}$  (red lines) S 2p components. The vertical colored bars are the theoretical IPs and shifted by  $-0.58$ ,  $-0.9$ , and  $-0.28$  eV for BBT, DBT and T, respectively.

This effect is known as *molecular-field splitting*, and it is usually small compared to the SO splitting. Since both the SO interaction and the molecular field splitting contribute to the shape and structure of the 2p core ionized/excited spectrum, both effects have been included in the computational scheme to obtain reliable simulations of the experimental spectra. Vibrational effects, which also contribute to the line-shape and broadening of the core ionized and excited spectra, have not been taken into account in the computational approach employed in this work. The assignment of the measured photoelectron bands is done with reference to Table 2 where the calculated IPs are reported together with the experimental BEs obtained from the fitting procedure.

As concerns the computed molecular field splitting along the series, it is found to be 120 and 140 meV for BBT and DBT, respectively (see Table 2), in rather good agreement with the experimental BE values derived from the fit (100 meV for BBT and 160 meV for DBT). Since a very similar experimental value is found for T (110 meV) it can be concluded that this reflects the comparable chemical and bonding environment of the S atom along the series.

Although in a high-resolution study by Giertz et al.<sup>24</sup> well resolved vibrational progressions in both C 1s and S 2p photoelectron spectra of thiophene were observed, the present experiment is not able to resolve the vibrational envelope of

**Table 2. Comparison between Theoretical TDDFT S 2p IPs and Experimental BEs Derived with the Fitting Procedure Described in the Text for BBT and DBT<sup>a</sup>**

edge	BBT			DBT			T <sup>c</sup>	
	IPs		BEs	IPs		BEs	IPs	BEs
	calculated	shifted <sup>a</sup>	single fit	calculated	shifted <sup>b</sup>	single fit	calculated	single fit
L <sub>II</sub>	171.53	170.95	170.82	171.47	170.57	170.47	171.63	171.14
L <sub>III</sub>	170.29	169.71	169.69	170.23	169.33	169.32	170.38	170.00
L <sub>III</sub>	170.17	169.59	169.59	170.09	169.19	169.16	170.27	169.89

<sup>a</sup>Calculated ionization potentials shifted by  $-0.58$  eV to match better the experimental peaks. <sup>b</sup>Calculated ionization potentials shifted by  $-0.90$  eV to match better the experimental peaks. <sup>c</sup>ref27. <sup>d</sup>Energies are in eV.

BBT and DBT. From Figure 3, we observe that the shape of the photoelectron spectra is similar along the series T-BBT-DBT: a more intensive photoelectron peak at lower BEs (169.89, 169.59, and 169.16 eV, respectively for T, BBT, and DBT) corresponds to the  $^2P_{3/2}$  state, and is separated by the SO splitting (1.18 eV) from the band relative to the  $^2P_{1/2}$  state. The spectra are characterized by asymmetric line profiles which suggest a vibrational contribution to the broadening. The fits reproduce very well the experimental data, also taking in account the molecular-field splitting of the S  $2p_{3/2}$  component.

From T to DBT, we observe a regular BE decrease of the corresponding photoelectron peaks. The shift to lower binding energies from T to BBT (of about 0.3 eV) reflects the increased shielding effect of the electronic charge density on Sulfur as a result of the presence of an annulated phenyl ring. The addition of a second phenyl ring to the BBT moiety further reduces the BE of DBT of about 0.37 eV. The experimental trend is qualitatively reproduced by theory although the shift among the calculated IPs along the series is underestimated by the latter. In particular, the KS eigenvalues overestimate the fitted BEs by as much as 0.58 eV for BBT and by about 0.90 eV for DBT. This overestimation is due to the too attractive character of the LB94 potential employed in the TDDFT calculations.  $\Delta$ KS calculation of IPs<sup>61</sup> as well as the use of relaxed orbitals could in principle recover the discrepancy with the observed trend. On the other hand, the agreement between the calculated and experimental SO splitting values is quite satisfactory for both BBT and DBT. Indeed, the computed SO splitting ( $\Delta_{3/2,1/2}$ ), expressed as the energy difference between the  $^2P_{1/2}$  state and the average of the  $^2P_{3/2}$  and  $^2P_{3/2}$  states, is equal to 1.3 eV for both molecules, in excellent agreement with the experimental one of about 1.20 eV, meaning that relativistic effects are rather well accounted for by the ZORA formalism.

**4.3. C K-edge NEXAFS.** In this section we present the NEXAFS measurements and calculations which give us a detailed understanding of the empty density of states of the studied molecules in the presence of the core hole. This implies that the energy position of the resonances observed in the spectra are affected by the Coulomb attraction between the core hole and the excited electron. Moreover also the character and the contribution of the different atomic sites to the virtual orbitals are modified by the presence of the core hole as understood by comparing the ground state (GS) LUMO and the TP LUMO orbitals of the BBT and DBT in the presence of a core hole on the C1 site, as examined in more details when discussing the NEXAFS results for BBT and DBT. We first discuss the C K-edge NEXAFS spectra of BBT and DBT in terms of the computed excitation energies and the transition

intensities (oscillator strengths) in order to assign the experimental features which are reported in Table 3 for BBT and in Table 4 for DBT. Then we will analyze the evolution of the spectral features with the increasing number of fused

**Table 3. Peak assignments for the C K-edge NEXAFS spectrum of BBT. Experimental energies are reported in the last column**

peak	site	calculated		assignment	experimental	
		$E$ shifted <sup>a</sup> (eV)	$f \times 10^2$ <sup>b</sup>		$E$ (eV)	peak
A	C3	285.09	2.73	$1\pi^*$ (LUMO)	284.93	1
	C7	285.17	2.27			
	C4	285.28	2.00		285.16	
	C5	285.29	2.57			
	C6	285.38	1.88		285.43	
	C8	285.50	2.97			
	C2	285.55	1.45			
	C6	285.63	1.07		$2\pi^*$ (LUMO+1)	
B	C1	285.93	2.25	$1\pi^*$	285.79	2
	C1	286.11	0.93	$2\pi^*$ (LUMO+1)		
C	C8	286.75	1.48	$\sigma^*(C-S)$	287.07	3
	C1	286.82	1.74			
	C4	286.92	1.65	$2\pi^*(C=C)$		
	C7	287.07	0.88			
	C2	287.09	1.63			
D	C8	287.52	0.42	$\sigma^*(C-H)$ - Rydberg mixed	287.61	–
	C6	287.62	0.46	$2\pi^*(C=C)$ - Rydberg mixed		
	C7	287.63	0.37	Rydberg		
	C3	287.67	0.32			
	C4	287.69	0.47			
E	C7	287.99	0.48	$\sigma^*(C-H)$ - Rydberg mixed	288.22	–
	C6	288.01	0.52			
	C8	288.06	0.41			
	C3	288.08	0.66			
	C4	288.12	0.50			
F	C8	288.50	1.01	$\pi^*(C=C)$ - Rydberg mixed	288.83	–
	C8	288.68	0.47	$\sigma^*(C-H)$ - Rydberg mixed		
G	C1	288.77	0.66	Rydberg		
	C7	289.16	0.28	Rydberg	289.29	–
	C6	289.17	0.35			
	C2	289.32	0.30			

<sup>a</sup>Calculated excitation energies shifted by  $-0.40$  eV to match the first experimental peak. <sup>b</sup>Only transitions with oscillator strength  $f \times 10^2 \geq 0.10$  are reported.

**Table 4. Peak Assignments for the C K-Edge NEXAFS Spectrum of DBT<sup>c</sup>**

peak	site	calculated		assignment	experimental	
		<i>E</i> shifted <sup>a</sup> (eV)	<i>f</i> × 10 <sup>2</sup> <sup>b</sup>		<i>E</i> (eV)	peak
A	C3	285.03	5.08	1 $\pi^*$ (LUMO)	284.95	1
	C6	285.09	4.45			
	C4	285.21	4.57		285.15	
	C5	285.32	2.47		285.36	
	C2	285.33	3.29			
	C5	285.49	3.39	2 $\pi^*$ (LUMO+1)	285.57	
B	C1	285.86	3.56	1 $\pi^*$	285.89	2
	C1	286.04	2.97	2 $\pi^*$		
C	C1	286.78	3.42	$\sigma^*$ (C–S)	287.15	3
D	C6	287.03	1.72	2 $\pi^*$		
	C2	287.10	2.41			
	C5	287.05	0.58	$\sigma^*$ (C–H)- Rydberg mixed		
	C4	287.24	1.00			
E	C5	287.50	0.89	$\pi^*$ (C=C)- Rydberg mixed	287.49	–
	C3	287.52	0.88			
	C6	287.55	0.68	Rydberg		
F	C5	287.91	1.00	Rydberg	288.07	–
	C6	287.95	0.99			
	C3	287.97	0.98	$\sigma^*$ (C–H)- Rydberg mixed		
G	C1	288.93	1.29	$\pi^*$ (C=C)- Rydberg mixed	288.91	–
	C5	289.00	1.22			
	C6	289.09	0.70			
	C2	289.14	0.69			

<sup>a</sup>Calculated excitation energies shifted by  $-0.40$  eV to match the first experimental peak. <sup>b</sup>Only transitions with  $f \times 10^2 \geq 0.10$  are reported <sup>c</sup>Experimental energies are reported in the last column.

phenyl rings in the moiety including also the T molecule, previously studied.<sup>26</sup> In the following discussion, the MOs involved in the most intense transitions of BBT and DBT will be considered and their composition analyzed in terms of Mulliken population of atomic orbitals (AOs) that are centered on each specific atom. The theoretical results for the C K-edge NEXAFS spectrum of the BBT molecule are listed in Table 3 and shown in Figure 4, together with the gas phase experimental data. The colored partial contributions of all nonequivalent *C<sub>i</sub>* atoms are also highlighted.

The experimental C K-edge NEXAFS spectrum of BBT (Figure 4) is characterized by a broad and structured resonance (peak 1) in the range 284.93–285.43 eV with a maximum of intensity centered at 285.16 eV and a second quite intensive feature (peak 2), on the high photon energy side, at 285.79 eV. The experimental data are shown along with the calculated spectra that help in identifying the contribution to the different peaks due to the transitions of the 1s electrons of the different *C<sub>i</sub>* atoms to the empty molecular states. The calculated total C K-edge spectrum is dominated by the low-energy peak (A) which is contributed by the C 1s  $\rightarrow$  LUMO (1 $\pi^*$ ) transitions from all the *C<sub>i</sub>* sites, with the exception of C1, whose C 1s  $\rightarrow$  LUMO transition is predicted to occur at higher energies (285.93 eV, structure B) also in agreement with the higher C 1s XPS IP of the C1 atom. The LUMO orbital of BBT referring to the TP calculation obtained from the core hole on C1 site is shown in the sketch of Figure 4. As

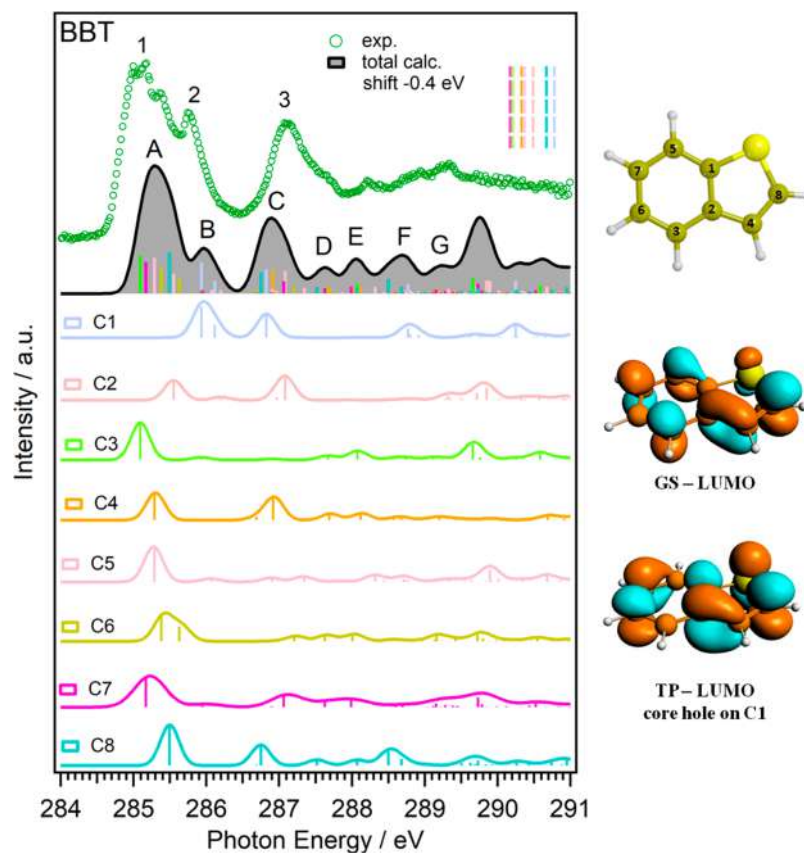
it can be easily seen, this TP - LUMO is delocalized on the entire molecule, highlighting the aromatic character of the molecule. Moreover, it also gets an increased contribution from the sulfur 3p atomic component passing from 6.90% (as calculated from the GS LUMO and shown in the middle sketch of Figure 4) to 17.24% (see bottom sketch of Figure 4). The C6 site contributes to peak A also with the transition to LUMO+1 (2 $\pi^*$ ) orbital, unlike the other *C<sub>i</sub>* sites of the phenyl ring whose LUMO+1 transitions do not acquire significant intensity. The C1(1s)  $\rightarrow$  LUMO and LUMO+1 (2 $\pi^*$ ) transitions, which are only separated by 0.2 eV, contribute to structure B. The much lower intensity of the latter with respect to the former agrees with the lower C1 2p AOs contributions to the LUMO+1 virtual orbital (around 8%) compared to the LUMO (around 19%). An inspection of the atomic site contributions, reported in Figure 4, suggests that, with the exception of C2 and C4 sites, the C 1s  $\rightarrow$  LUMO transition carries the largest intensity in line with the larger *C<sub>i</sub>* 2p AOs contribution to the LUMO with respect to the higher virtual MOs.

The atomic site dependence of computed *C<sub>i</sub>*(1s)  $\rightarrow$  LUMO excitation energies mostly follows the trend already observed for the *C<sub>i</sub>*(1s) XPS BEs, with the exception of C6, whose transition is pushed at relatively higher energy leading us to conclude that final state effects (e.g., relaxation effects) are site-specific and assume a certain importance for C6. As already reported by Zhang et al.<sup>65</sup> for DPTA molecule (a substituted Triphenylamine with a thiophene ring) the core hole site can have a significant effect on the transition energy observed in NEXAFS.

To Peak 3 (at 287.07 eV in the experimental spectrum), which is the second most intense structure of both theoretical and experimental spectra, many transitions contribute, of which the most intense involve the  $\pi^*$ (C=C) and the  $\sigma^*$ (C–S) virtual antibonding orbitals (see Table 3). Transitions to the delocalized  $\pi^*$  involve *C<sub>i</sub>* atoms not directly bonded to the heteroatom, while transitions to the more localized  $\sigma^*$ (C–S) virtual MOs involve C8 and C1 sites (at 286.75 and 286.82 eV by theory, respectively). These virtual MOs present also significant contributions from the S 2p and 3d AOs. The higher energy structures, labeled D, E, F and G, derive their intensity from a manifold of transitions where the final MOs have either mixed valence/Rydberg or pure Rydberg character. The valence character of the more intense transitions involves both  $\pi^*$ (C=C) and  $\sigma^*$ (C–H) characters, with the exception of structure E, where all transitions with appreciable intensity involve final states with mixed  $\sigma^*$ (C–H)/Rydberg character. As expected, transitions to diffuse virtual MOs are characterized by a lower intensity. Just below the ionization threshold (peak G, at around 289.20 eV and the following broader and intense feature), the spectrum is characterized by a series of transitions to Rydberg states, converging toward the ionization threshold.

Every spectral band in the theoretical spectrum has a counterpart in the experimental one so that the agreement is very good, although the energy separation between features B and C–D is slightly underestimated by the DFT-TP methodology used in the present work. Nevertheless, the general good agreement permits a conclusive assignment of the experimental spectrum, as indicated in Table 3.

The good agreement between the experimental and the theoretical NEXAFS spectra are also found for the hereafter reported DBT molecule in Figure 5 and Table 4. The



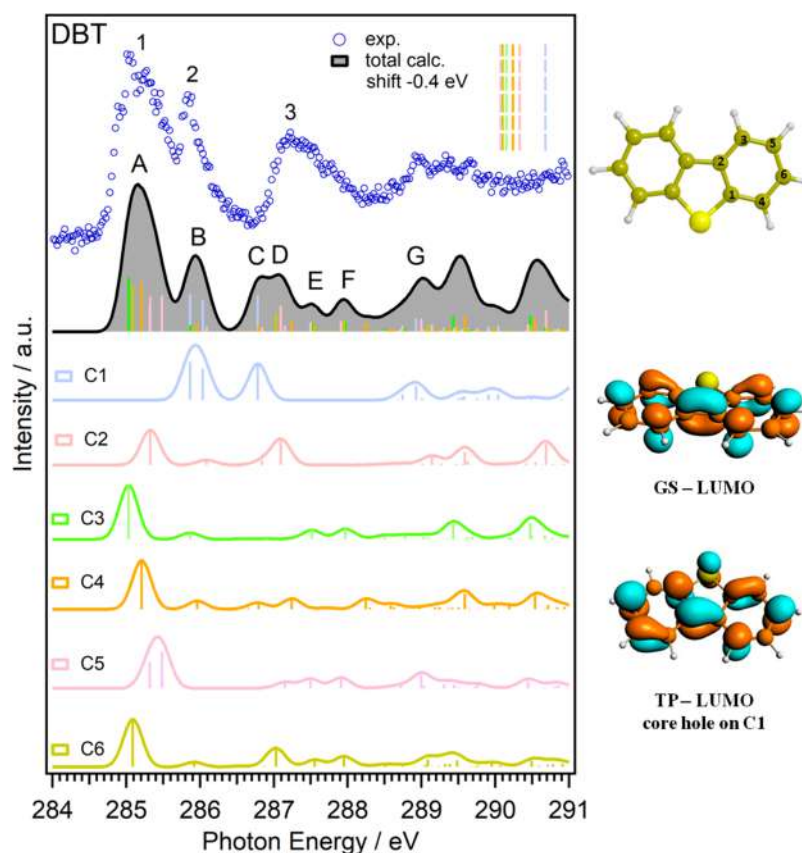
**Figure 4.** Left side: C K-edge NEXAFS spectrum of BBT. The experimental spectrum (green circles) is shown together with the total theoretical line shape (solid black line) and the partial Ci contributions (colored vertical lines). Partial Ci contributions convoluted by Gaussian profiles with  $fwhm = 0.3$  eV are presented as colored spectra in the bottom panels. The  $\Delta KS$  C 1s ionization thresholds are also shown (colored vertical dashed bars). The calculated peaks have been shifted by  $-0.4$  eV in order to match the experimental results. Right side: molecular structure of BBT with labels for the different Ci atoms (top); the GS-LUMO (middle) and the TP-LUMO with a core hole on C1 (bottom).

experimental C K-edge NEXAFS spectrum of DBT shows a broader and intense low-energy peak (1) in the range 284.95–285.57 eV and centered at about 285.15 eV partially overlapped with a less intense feature (2, at 285.89 eV) which acquires gradually more intensity with the progressively thiophene annulation with one and two fused phenyl rings respectively in BBT and DBT molecules. These two structures are then followed by a double-peaked broad spectral feature (peak 3 at around 287.15 eV), while less resolved structures (E, F, and G as calculated peaks) fall in the higher excitation energy region before the ionization threshold.

Peak A is assigned to the C 1s  $\rightarrow$  LUMO ( $\pi^*$ ) transitions from all the nonequivalent C sites, with the exception of C1. Indeed, the largest C 1s  $\rightarrow$  LUMO ( $\pi^*$ ) excitation energy is predicted for the C1 site which is less screened being directly bonded to the S atom, as already observed for the BBT molecule, and in agreement with the energy position of the calculated IPs. Peak B is contributed to by both C1(1s)  $\rightarrow$  LUMO ( $\pi^*$ ) and C1(1s)  $\rightarrow$  LUMO+1 ( $2\pi^*$ ), which are characterized, at variance with the BBT case, by a comparable intensity reflecting a similar C 2p AOs contribution to both virtual MOs. The S 3p contribution to the LUMO is even more pronounced when passing from GS- to TP-LUMO after creating a core hole on C1 site (compare middle and bottom sketches of Figure 4 and Figure 5, respectively). Indeed, from calculation an increased contribution from the sulfur 3p atomic component (7.42%) for TP-LUMO has been found compared to the GS-LUMO for which the sulfur 3p contribution is of

only 1.12%. Moreover, as before described for BBT, also in the case of DBT, the TP-LUMO clearly shows the aromatic delocalization on the whole system proving the aromatic character of the DBT molecule. Another interesting aspect of the theoretical spectrum can be obtained by analyzing the site-resolved partial contributions: for the core-excitation of C5 the transition C 1s  $\rightarrow$  LUMO+1 is more intense than the C 1s  $\rightarrow$  LUMO, with the higher weights (around 16%) of C5 (2p) AOs contributions to the former MO compared to the latter (around 10%). The importance of electronic relaxation following the core-hole formation in the case of C5 is also evident by inspecting and comparing the site-dependence of C 1s  $\rightarrow$  LUMO excitation energy with the trend in calculated IPs: the C5 site has the lowest IP value while its  $1\pi^*$  transition is the highest, with the exception of the C1 and C2 heterocyclic sites. This fact indicates that final-state effects are quite important in this case, and are in general strongly site-specific.<sup>65</sup>

Peak (C) (at 286.78 eV photon energy of the calculated spectrum) is mainly contributed by the C1 1s transition toward a  $\sigma^*(C-S)$  state, while the following three structures (D, E, F) are characterized by a superposition of core excitations to virtual MOs of valence and Rydberg character of the C atoms not directly bonded to the S atom. The progressively decreasing intensity of transitions to virtual orbitals with  $\pi^*$  character with increasing excitation energy reflects the general reduction of the C 2p valence character of the higher energy virtual MOs. In the same spectral region, transitions of lower-



**Figure 5.** Left side: C K-edge NEXAFS spectrum of DBT. Experimental spectrum (blue circles) is shown together with the total theoretical line shape (solid black line) and the partial C<sub>i</sub> contributions (colored vertical lines). Partial C<sub>i</sub> contributions convoluted by Gaussian profiles with fwhm = 0.3 eV are presented as colored spectra in the bottom panels. The  $\Delta$ KS C 1s ionization thresholds are also shown (colored vertical dashed bars). The calculated peaks have been shifted by  $-0.4$  eV in order to match the experimental results. Right side: molecular structure of DBT with atoms labels (top); the GS-LUMO (middle) and the TP-LUMO with a core hole on C1 (bottom).

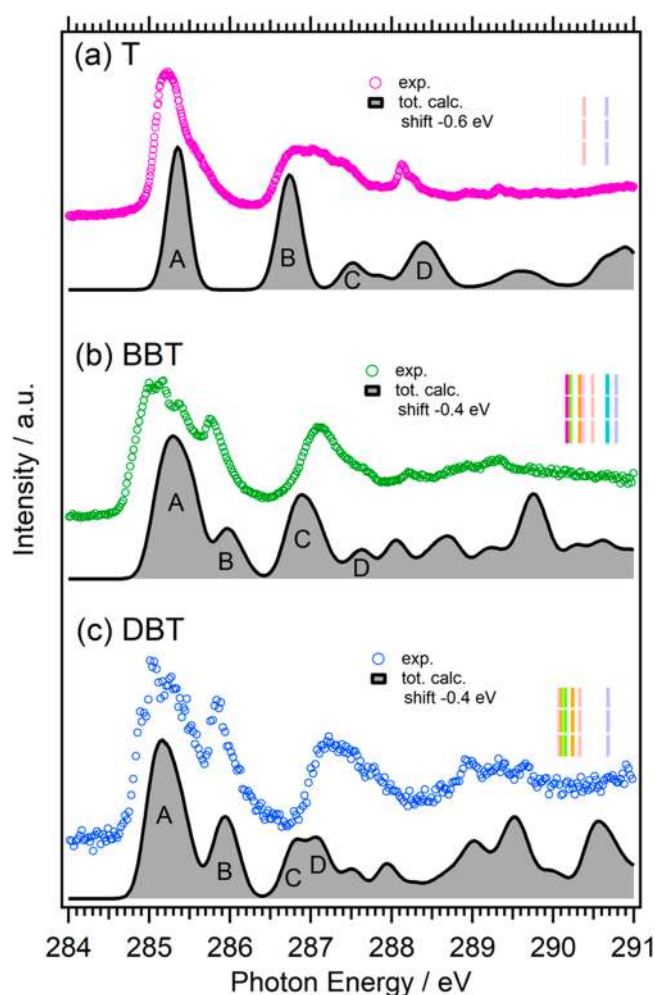
intensity to MOs with mixed valence  $\sigma^*(\text{C-H})$ /Rydberg character and high weights of C (3p,3d) AOs are also visible. Finally, peak G (at around 289.00 eV) is ascribed to transitions to Rydberg states with some  $\pi^*(\text{C=C})$  contributions; the mixed valence–Rydberg nature of these transitions is confirmed by their significant intensity.

We finally address the evolution of the experimental NEXAFS spectral features along the series T, BBT, and DBT molecules (Figure 6), i.e., for increasing number of fused phenyl rings in the moiety. The more pronounced spectral variations are visible in the lower energy region of the spectra, where a single peak with a high-energy tail in T acquires a pronounced substructure in BBT (peaks B), and finally splits into two well resolved peaks in DBT. It is also important to note that the position of the first peak is shifted to lower energies (by about 0.50 eV) in going from T to BBT, while it does not change further from BBT to DBT: this trend could be explained by the stabilizing effect due to aromaticity, which is more pronounced in passing from one to two fused rings than with the introduction of a third ring. This is also reflected by the different BEs observed in the C 1s XP spectra of these three molecules. Earlier calculations<sup>26</sup> attributed the asymmetric shape of the first NEXAFS band in T to its vibrational envelope, which can be also recognized in the spectra of BBT and DBT. Purely electronic initial state and relaxation effects are instead responsible of the appearance and evolution of structure B in going from BBT to DBT. The second spectral structure visible in T can be related to structures C and C–D

in BBT and DBT, respectively, while the higher energy structures reveal a more complex nature, arising from transitions to mixed valence–Rydberg and Rydberg states. The number of overlapping transitions increases in the high energy region of the spectra, preventing a strict correspondence along the series. To conclude, the agreement between theory and experiment is globally satisfactory, the only exception being the energy separation between structures B and C in the two derivatives of thiophene, somewhat underestimated by the DFT-TP approach compared to the experiments.

**4.4. S L<sub>II,III</sub>-edge NEXAFS.** In Figure 7 the S L<sub>II,III</sub>-edge NEXAFS spectra of BBT (top panel) and DBT (bottom panel) are shown together with the theoretical calculated results for identifying the contribution to the spectra considering the SO splitting of the S 2p orbital. The experimental spectrum of BBT has three distinct resonances at about 165.13 eV, a more intensive peak at 166.93 eV with a clear shoulder at the low photon energy side (at about 166.22 eV) and then the third feature at 168.32 eV. The spectrum of DBT has instead four resolved resonance peaks centered at about at 165.15, 166.24, 166.90, and 167.55–168.15 eV.

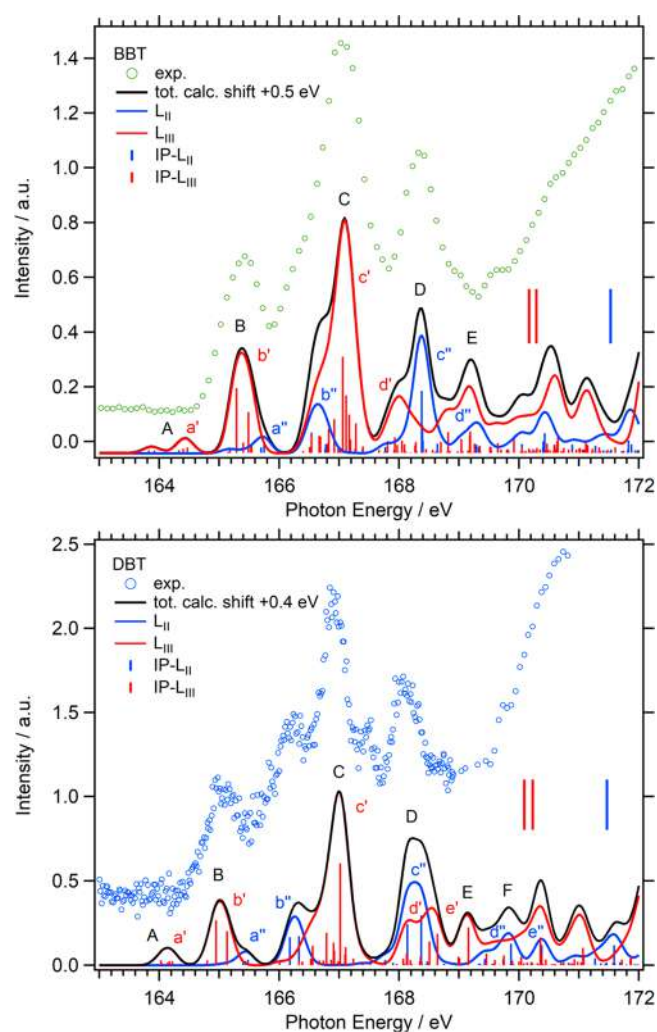
In the S 2p core excitation spectrum, the SO splitting of the core hole in the 2p<sub>3/2</sub> and 2p<sub>1/2</sub> components leads to distinct spectral features converging to L<sub>III</sub> and L<sub>II</sub> ionization thresholds. The spectral assignments are therefore relatively difficult, since the SO energy splitting is quite small (1.30 eV) and, as a consequence, there is a partial overlap of the two



**Figure 6.** Comparison of C K-edge NEXAFS spectra of: (a) T, (b) BBT, and (c) DBT. Experimental spectra (colored circles) are shown together with the total theoretical line-shape (solid black line) obtained by Gaussian profiles with  $\text{fwhm} = 0.3$  eV. The  $\Delta\text{KS C } 1s$  ionization thresholds are also shown (colored vertical dashed bars). The calculated peaks have been shifted by  $-0.4$  (BBT and DBT) and  $-0.6$  eV (T) in order to match the experimental results.

manifolds of excited states. Moreover, the assignment is further complicated by the intensity redistribution among the final excited states due to the configuration mixing included in the TDDFT scheme. However, despite the needed complex approach the theoretical results are in excellent agreement with the experimental spectra, giving a reliable description of the different transitions contributing to the observed intensities.

In the analysis of the spectral features it is convenient to first perform a deconvolution of the calculated S 2p line spectra into the two manifolds of excited states converging to the  $L_{\text{III}}$  and  $L_{\text{II}}$  edges (see Figure 7). The deconvolution is obtained by multiplying, for each transition, the computed oscillator strength by the weight of the contributing  $1h-1p$  configurations corresponding to excitations from the S  $2p_{3/2}$  and S  $2p_{1/2}$  levels, respectively. To facilitate the comparison with the experiment, the bars spectra are further broadened by using Gaussian functions with  $\text{fwhm}$  of 0.30 eV. The assignment of the most intense transitions to the character of the final MOs of the dominant  $1h1p$  excitations contributing to each final excited state are collected in Table 5 and Table 6 for BBT and



**Figure 7.** S  $L_{\text{II,III}}$ -edge NEXAFS spectra of BBT (top panel) and DBT (bottom panel): experimental data (colored circles), calculated TDDFT results (black solid line). Also shown is the deconvolution of the calculated S 2p spectra into the two manifold of excited states converging to the  $L_{\text{III}}$  (red solid line and vertical red bars) and  $L_{\text{II}}$  (blue solid line and vertical blue bars) edges. The energy scale of the calculated data has been shifted by  $+0.5$  and  $+0.4$  eV, respectively for BBT and DBT in order to match the first experimental peak. The DFT S 2p ionization thresholds are also shown (blue and red vertical bars) within each spectrum.

DBT, respectively. The calculated oscillator strength for S 2p core excitations maps the sulfur s and d contents in the final MOs.

In BBT, the S 2p orbitals in the relativistic SO description correspond to the  $12a_{1/2}$ ,  $13a_{1/2}$  ( $L_{\text{III}}$  components, split by the molecular field), and  $11a_{1/2}$  ( $L_{\text{II}}$  component) spinors. The lowest energy transitions of the  $2p_{3/2}$  electrons are toward the LUMO ( $7a''$  MO) and LUMO+1 ( $8a''$  MO) virtual orbitals; both of them exhibit an antibonding  $\pi^*$  character delocalized on the entire moiety. The calculated low intensity of the implicated transitions reflects the small contribution of S  $nd$  orbitals to the final MOs, with respect to the high intensity carried by the transitions toward these MOs in the C K-edge NEXAFS spectrum. The peak A is barely visible in the experimental profile, suggesting a possible overestimation of the S  $nd$  contributions to these MOs.

Table 5. Peak Assignments for the S L<sub>II,III</sub>-Edge NEXAFS Spectrum of BBT<sup>c</sup>

peak/subpeak	calculated			EDGE	assignment	experimental		
	E (eV)	E shifted <sup>a</sup> (eV)	$f \times 10^2$ <sup>b</sup>			E (eV)		
A/a'	163.28	163.78	0.043	L <sub>III</sub>	7a''/1π* (LUMO)	–		
	163.42	163.92	0.015					
	163.90	164.40	0.026		8a''/2π* (LUMO+1)			
	163.97	164.47	0.032					
B/a''	165.20	165.70	0.034	L <sub>II</sub>	8a''	–		
	165.25	165.75	0.069					
B/b'	164.68	165.18	0.035	L <sub>III</sub>	30a'/σ*(C–S) (LUMO+2)	165.13		
	164.79	165.29	0.400					
	164.90	165.40	0.064					
	164.99	165.49	0.271		9a''/π* (LUMO+3)			
	165.04	165.54	0.060					
	165.90	166.40	0.011		31a'/σ*(C–H)-Rydberg mixed			
C/b''	166.04	166.54	0.238	L <sub>II</sub>	30a'	~166.22		
	166.17	166.67	0.204					
	166.27	166.77	0.026		9a''			
	166.28	166.78	0.058					
C/c'	166.19	166.69	0.149	L <sub>III</sub>	32a'/σ*(C–H)-Rydberg mixed	166.93		
	166.33	166.83	0.164					
	166.42	166.92	0.206		34a'/σ*(C–H)-Rydberg mixed			
	166.56	167.06	0.588		33a'/σ*(C–H)-Rydberg mixed			
	166.62	167.12	0.352		33a'			
	166.67	167.17	0.236					
	166.78	167.28	0.180		35a'/Rydberg nsS			
	167.27	167.77	0.038		L <sub>II</sub>		31a'	168.32
	167.88	168.38	0.385				33a'	
	167.89	168.39	0.179					
D/d'	167.90	168.40	0.078	L <sub>III</sub>				
	168.14	168.64	0.057		35a'			
	167.17	167.67	0.019		36a'/σ*(C–H)-Rydberg mixed			
	167.31	167.81	0.075					
	167.43	167.93	0.095		10a''/π*			
	167.47	167.97	0.050		37a'/σ*(C–H)-Rydberg mixed			
	167.55	168.05	0.076		11a''/π*			
	167.78	168.28	0.120		11a''			
	167.95	168.45	0.024		38a'/σ*(C–H)-Rydberg mixed			
	168.06	168.56	0.018					
E/d''	168.20	168.70	0.068	L <sub>II</sub>	39a'/σ*(C–H)-Rydberg mixed	~169.03		
	168.54	169.04	0.107		36a'			
	168.83	169.33	0.038		37a'			
	169.03	169.53	0.038		42a'/Rydberg ndS			

<sup>a</sup>Calculated excitation energies shifted by +0.50 eV to match the energy position of the first experimental peak. <sup>b</sup>Only the main transitions are reported. <sup>c</sup>Experimental energies are reported in the last column.

The first intense peak of BBT (peak B, at around 165.13 eV in the experimental spectrum) mainly arises from the S 2p<sub>3/2</sub> excitations (subpeak b', in red). The two most intense lines (at 165.29 and 165.49 eV, see Table 5) correspond to transitions to the LUMO+2 (30a') of σ\*(C–S) character and LUMO+3 (9a'', π\*) MOs. Their relatively high intensity is due to the significant contribution of S3d orbitals to these MOs. Peak B is also partially contributed by excitations with lower intensity toward the LUMO+1 and converging to the L<sub>II</sub> edge (subpeak a'', in blue). The energy separation between the SO partner transitions giving rise to the in a' and a'' subpeaks is very close (1.23 eV) to the experimental SO splitting (1.20 eV).

Structures A and B are then followed by the predominant spectral feature (peak C, at around 166.93 eV in the experimental spectrum) which is mainly contributed by transitions converging to the L<sub>III</sub> edge (subpeak c') and

involving final MOs with a mixed valence–Rydberg nature. The main intense transitions are toward virtual orbitals with partial valence antibonding character and high weights of S diffuse AOs. The pronounced shoulder at lower energy is contributed by subpeak b'' (in blue) which represents the L<sub>II</sub> partner transitions of the L<sub>III</sub> component relative to b' peak.

The SO partner of subpeak c' (subpeak c'') significantly contributes to peak D, at around 168.32 eV in the experimental spectrum; the relatively high intensity of the main transitions reflects the significant ns and nd AOs contributions from the S atom to valence σ\*(C–H) final orbitals. The lower energy shoulder of peak D derives its intensity from a manifold of transitions (subpeak d') to virtual states of mixed valence/Rydberg or pure Rydberg character. Features at higher excitation energy (peak E) are a result of a large number of



Table 6. Peak Assignments for the S 2p NEXAFS Spectrum of DBT<sup>c</sup>

peak/subpeak	calculated			EDGE	assignment	experimental
	E (eV)	E shifted <sup>a</sup> (eV)	$f \times 10^2$			E (eV)
A/a'	163.63	164.03	0.044	L <sub>III</sub>	4a <sub>2</sub> /2π* + nd S (LUMO+1)	–
	163.74	164.14	0.023		7b <sub>1</sub> /3π* + nd S (LUMO+2)	
	163.77	164.17	0.033		4a <sub>2</sub>	
	163.82	164.42	0.049		7b <sub>1</sub>	
B/a''	165.00	165.40	0.023	L <sub>II</sub>	4a <sub>2</sub>	–
	165.04	165.44	0.036		7b <sub>1</sub>	
	165.08	165.48	0.046			
B/b'	164.40	164.80	0.040	L <sub>III</sub>	19b <sub>2</sub> /σ*(C–S)-Rydberg mixed (LUMO+3)	165.15
	164.55	164.95	0.381			
	164.73	165.13	0.306			
C/b''	165.78	166.18	0.240	L <sub>II</sub>	19b <sub>2</sub>	166.24
	165.93	166.33	0.293			
C/c'	166.16	166.56	0.174	L <sub>III</sub>	20b <sub>2</sub> /σ*(C–S)-Rydberg mixed	169.90
	166.39	166.79	0.273			
	166.51	166.91	0.192			
	166.62	167.02	0.864		26a <sub>1</sub> /mixed valence–Rydberg	
	166.71	167.01	0.155			
D/c''	167.74	168.14	0.395	L <sub>II</sub>	25a <sub>1</sub>	167.55
	167.97	168.37	0.436		26a <sub>1</sub>	
D/d'	167.73	168.13	0.125	L <sub>III</sub>	22b <sub>2</sub> /σ*(C–H)-Rydberg mixed	168.15
	167.96	168.36	0.109			
	168.11	168.51	0.205			
	168.24	168.64	0.266		29a <sub>1</sub> /σ*(C–H)-Rydberg mixed	
E/e'	168.76	169.16	0.328	L <sub>III</sub>	30a <sub>1</sub> / 3d S	169.45
F/d''	169.06	169.46	0.111	L <sub>II</sub>	31a <sub>1</sub> /σ*(C–H)-Rydberg mixed	~169.73
	169.19	169.59	0.017			
	169.35	169.75	0.116		8a <sub>2</sub> /Rydberg ndS	
	169.47	169.87	0.183		32a <sub>1</sub> /Rydberg nsS	
	169.49	169.89	0.042			

<sup>a</sup>Calculated excitation energies shifted by +0.40 eV to match the energy position of the first experimental peak. <sup>b</sup>Only the main transitions are reported. <sup>c</sup>Experimental energies are reported in the last column.

weak transitions to Rydberg states, converging to both the L<sub>II</sub> and L<sub>III</sub> thresholds.

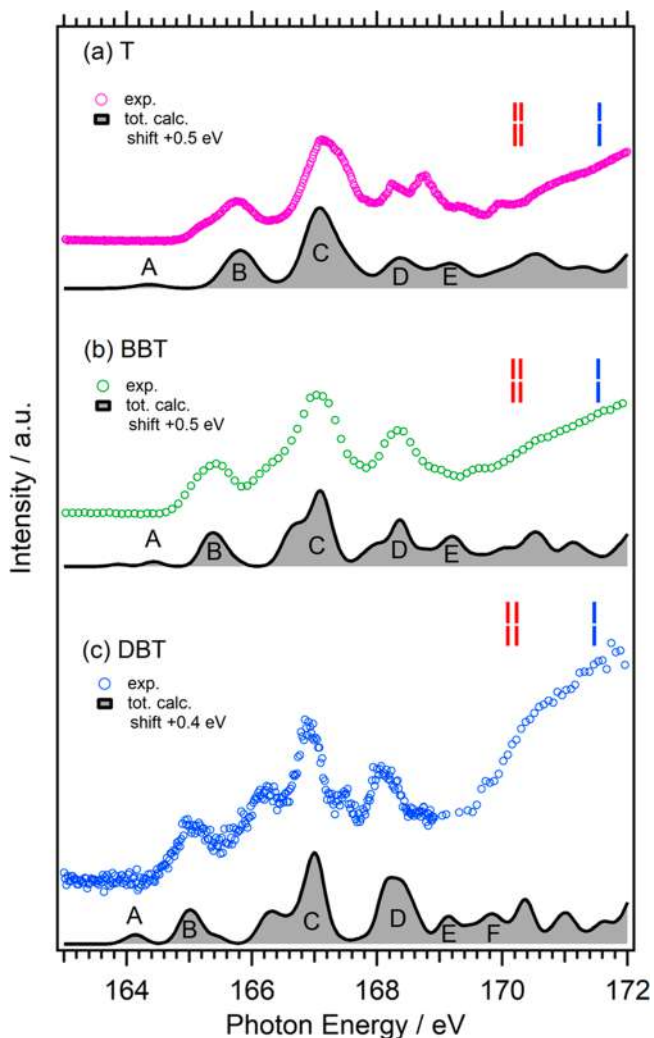
The S 2p orbitals of DBT in the relativistic SO description correspond to the 16e<sub>1/2</sub>, 17e<sub>1/2</sub> (L<sub>III</sub> components, split by the molecular field), and 15e<sub>1/2</sub> (L<sub>II</sub> component) spinors. The assignment of the spectral features is reported in Table 6. The theoretical NEXAFS profile of DBT shares some similarities with the corresponding spectrum of the BBT molecule (compare top and bottom panels of Figure 7). The S 2p → LUMO transition is not visible in the calculated (and neither in the experimental) spectrum due to the negligible contribution of the S ns,d AOs to this MO. A low-lying weak spectral feature (peak A) arises from transitions originating from the S 2p<sub>3/2</sub> levels to the LUMO+1 and LUMO+2 virtual orbitals, both of strong π\* valence antibonding character. The low intensity of these transitions reflects the very small S nd AOs contributions to the virtual MOs. The corresponding transitions starting from the S 2p<sub>1/2</sub> level converging to the L<sub>II</sub> edge fall at around 165.4 eV (subpeak a'', in blue) and contribute to the high-energy shoulder of band B. As in the BBT case, the closeness between the energy separation of the a' and a'' subpeaks (1.18 eV) and the experimental SO splitting (1.20 eV, previously shown in S 2p XPS), suggests that relaxation (final state) effects are not important in this energy region; moreover the presence of peak A is barely visible in the experimental profile, pointing toward

an overestimation of the S nd contribution to the involved MOs.

As shown in Table 6, peak B is mainly contributed by subpeak b', which arises from L<sub>III</sub> transitions toward the LUMO+3 (19b<sub>2</sub>) at about 165 eV, which has a σ\*(C–S) valence character and which exhibits significant S diffuse AOs. Its higher intensity compared to peak A is consistent with the larger Snd AOs contributions to the LUMO+3 MO compared to the LUMO+1 and LUMO+2. Structure B is followed by a broader and stronger band, C, which dominates the NEXAFS spectrum as in BBT, and is characterized by a low-energy shoulder which appears more pronounced than in the BBT spectrum. The shoulder originates from the excitations of the 2p<sub>1/2</sub> electrons toward the LUMO+3 (subpeak b''). A manifold of transitions converging to the L<sub>III</sub> edge are responsible for the high intensity of structure C (subpeak c'); in particular, the more intense transition at 167.02 eV involves a virtual MO with mixed valence/Rydberg character and mainly contributed by s, p, and d AO functions of S atom. Its SO partner feature (subpeak c'', blue) contributes to peak D, together with transitions to virtual orbitals which have a mixed valence–Rydberg nature and converge to the L<sub>III</sub> edge (subpeak d', red). The energy separation between peaks C and D is however slightly overestimated by the theory. Structures E and F originates from transitions to mixed valence–Rydberg and pure Rydberg states.

The SO partner of subpeak e' (subpeak e'') is centered at around 170 eV, just above the  $L_{III}$  thresholds; therefore, it is not included in Table 6.

We conclude this section by analyzing the evolution of experimental and theoretical S 2p NEXAFS spectra along the series T-BBT-DBT (see Figure 8); the assignment of the S 2p



**Figure 8.** Comparison of S  $L_{II,III}$ -edge NEXAFS spectra of (a) T, (b) BBT, and (c) DBT. Experimental spectrum (colored circles) are shown together with the total calculated line-shape (solid black line) obtained by Gaussian profiles with  $fwhm = 0.3$  eV. The calculated ionization thresholds are also shown (red and blue vertical dashed bars). The energy scale is the one coming from experiments and the calculated peaks have been shifted by +0.5 eV (T and BBT) and +0.4 eV (DBT) in order to match the experimental results.

NEXAFS spectrum of T is reported in ref 27. All three spectra are dominated by band C, located at around 166–167 eV: it appears as a single peak with a strong asymmetry on the high energy side in the T spectrum, and partially overlapped with band D, it acquires a pronounced shoulder in BBT and shows a clear double-peak profile in DBT. The nature of the peaks is however conserved along the series, being contributed by  $L_{II} \rightarrow \sigma^*(C-S)$  transitions and  $L_{III}$  transitions of mixed valence–Rydberg character, similarly to what already found when examining the evolution of the S 2p spectra in a series of short chain oligothiophenes.<sup>27</sup> Since generally the main features are

conserved along the series, both as concerns the energy positions and the nature of the more intense transitions, this suggests that the electronic and geometrical environment of the S atom in the three molecules is little affected by the increasing number of fused phenyl rings. The relatively simple structure of the spectral profiles along the series reflects the localized nature of the virtual states involved in the S 2p core excitation process.

## 5. CONCLUSIONS

In this work we present a combined experimental and theoretical study of the electronic structure of benzo[*b*]thiophene (BBT) and dibenzothiophene (DBT) in the gas phase, by combining near-edge x-ray-absorption fine-structure (NEXAFS) and X-ray photoelectron spectroscopy (XPS) spectroscopies and theoretical calculations based on density functional theory (DFT) and its time-dependent generalization (TDDFT) in the linear response regime. The overall very good agreement between theoretical results and experimental data observed permitted an unambiguous assignment of the XPS and NEXAFS spectral features. A comparison of the C 1s XP spectra of the three molecules reveals the presence of an inductive effect of the S atom, and singles out effects due to the progressive increase of charge density for aromatic conjugation along the series. The calculation of the S 2p XPS allowed an analysis of binding energies in terms of both SO and molecular-field splitting, of the degenerate  $2p_{3/2}$  level, both of which are relatively conserved in the series, and due to the comparable chemical and bonding environment in the series.

As concerns the measured C K-edge NEXAFS spectra, TP-DFT calculations provide a good account of electron correlation phenomena related to both the aromatic stabilization along the series T-BBT-DBT as well as inductive and relaxation effects due to the presence of the sulfur atom. More specifically, the evolution of the NEXAFS spectral features along the series T-BBT-DBT points out that the stabilizing effect due to increased aromaticity is more pronounced in going from T to BBT compared to the introduction of a second annulated phenyl ring, analogously to what found in a related study on short-chain oligothiophenes.<sup>26</sup>

The analysis and spectral attribution of the experimental S  $L_{II,III}$ -edge NEXAFS data, quite demanding from a theoretical standpoint, due to the presence of overlapping manifolds of excited states converging to the  $L_{II,III}$  edges, points toward a much lower dependence of the observed spectral features on the degree of annulation, since the nature and spectral positions of the most intense NEXAFS spectral features are conserved along the series and reflects the localized nature of the virtual states involved in the S 2p core-excitation process. Thus, unlike the C 1s NEXAFS spectra, the analysis of the S 2p NEXAFS intensity is not able to capture in detail the effects of increased aromaticity along the series.

The comprehensive characterization of these molecules is of fundamental importance for being able to predict the effect of combining such moieties into more complex molecular systems. The benzo-annulated thiophene systems are in fact promising candidates for organic molecular electronics providing a high degree of  $\pi$ -conjugation due to the rigidified, planar structure which intrinsically affords smaller HOMO–LUMO gaps for the resulting material with respect to their not benzo-annulated building blocks. It will be interesting to characterize the electronic structure of such benzo-annulated

thiophene units implemented in promising light-harvesting molecules<sup>66</sup> or as  $\pi$ -conjugated core for phosphorescent light emitting units.<sup>17,18</sup>

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.8b08333.

Comparison between theoretical and experimental geometrical parameters for T (Tables S1) and for DBT (Table S2) and  $\Delta$ KS C 1s IPs of BBT and DBT computed by using the hybrid B3LYP xc potential (Table S3) (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

\*(A.G.) E-mail: [ambra.guarnaccio@pz.ism.cnr.it](mailto:ambra.guarnaccio@pz.ism.cnr.it).

\*(G.F.) E-mail: [fronzoni@units.it](mailto:fronzoni@units.it).

### ORCID

D. Toffoli: 0000-0002-8225-6119

T. Zhang: 0000-0001-8739-7773

F. Johansson: 0000-0002-6471-1093

C. Puglia: 0000-0001-6840-1570

M. Stener: 0000-0003-3700-7903

G. Fronzoni: 0000-0002-5722-2355

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors thank the Carl Trygger Foundation for financial support and for making available the VG-Scienta SES-200 photoelectron analyzer at the Gas Phase beamline, Elettra, Italy. Computational research has been supported by Finanziamento per ricerca di ateneo, FRA 2015 and FRA 2016 of the Università degli Studi di Trieste. The authors acknowledge the EU CERIC–ERIC Consortium for the access to experimental facilities and financial support. We thank Mr. G. Bortoletto and Mr. C. Pedersini of the User Support Lab at Elettra.

## ■ REFERENCES

- (1) Xiang, D.; Wang, X.; Jia, C.; Lee, T.; Guo, X. Molecular-Scale Electronics: From Concept to Function. *Chem. Rev.* **2016**, *116*, 4318–4440.
- (2) Park, S.; Wang, G.; Cho, B.; Kim, Y.; Song, S.; Ji, Y.; Yoon, M.-H.; Lee, T. Flexible Molecular-Scale Electronic Devices. *Nat. Nanotechnol.* **2012**, *7*, 438–442.
- (3) Roncali, J.; Leriche, P.; Blanchard, P. Molecular Materials for Organic Photovoltaics: Small is Beautiful. *Adv. Mater.* **2014**, *26*, 3821–3838.
- (4) Collins, S. D.; Ran, N. A.; Heiber, M. C.; Nguyen, T. Q. Small is Powerful: Recent Progress in Solution Processed Small Molecule Solar Cells. *Adv. Energy Mater.* **2017**, *7*, 1602242.
- (5) Shirota, Y.; Kageyama, H. Small Molecular Weight Materials for (Opto)electronic Applications: Overview. *Handbook of Organic Materials for Optical and (Opto)electronic Devices* **2013**, 3–28.
- (6) Baert, F.; Cabanetos, C.; Leliège, A.; Kirchner, E.; Segut, O.; Alévêque, O.; Allain, M.; Seo, G.; Jung, S.; Tondelier, D.; et al. Bridged Low Band Gap A–D–A Quaterthiophene as Efficient Donor for Organic Solar Cells. *J. Mater. Chem. C* **2015**, *3*, 390–398.
- (7) Bouzzine, S. M.; Makayssi, A.; Hamidi, M.; Bouachrine, M. Bridging Effect on Structural and Optoelectronic Properties of Oligothiophene. *J. Mol. Struct.: THEOCHEM* **2008**, *851*, 254–262.

(8) Haid, St.; Marszalek, M.; Mishra, A.; Wielopolski, M.; Teuscher, J.; Moser, J.-E.; Humphry-Baker, R.; Zakeeruddin, S. M.; Grätzel, M.; Bäuerle, P. Significant Improvement of Dye-Sensitized Solar Cell Performance by Small Structural Modification in  $\pi$ -Conjugated Donor–Acceptor Dyes. *Adv. Funct. Mater.* **2012**, *22*, 1291–1302.

(9) Zhang, F.; Wu, D.; Xu, Y.; Feng, X. Thiophene-Based Conjugated Oligomers for Organic Solar Cells. *J. Mater. Chem.* **2011**, *21*, 17590–17600.

(10) Turkoglu, G.; Cinar, M. E.; Ozturk, T. Thiophene-Based Organic Semiconductors. *Top. Curr. Chem.* **2017**, *375*, 1–45.

(11) Fichou, D.; Ziegler, C. Structure and Properties of Oligothiophenes in the Solid State: Single Crystals and Thin Films. *Handbook of Oligo- and Polythiophenes* **2007**, 183–282.

(12) Duan, C.; Gao, K.; Colberts, F. J. M.; Liu, F.; Meskers, S. C. J.; Wienk, M. M.; Janssen, R. A. J. Thiophene Rings Improve the Device Performance of Conjugated Polymers in Polymer Solar Cells with Thick Active Layers. *Adv. Energy Mater.* **2017**, *7*, 1700519.

(13) Zhang, G.; Bai, Y.; Li, R.; Shi, D.; Wenger, S.; Zakeeruddin, S. M.; Grätzel, M.; Wang, P. Employ a Bisthienothiophene Linker to Construct an Organic Chromophore for Efficient and Stable Dyesensitized Solar Cells. *Energy Environ. Sci.* **2009**, *2*, 92–95.

(14) Tian, Z.; Huang, M.; Zhao, B.; Huang, H.; Feng, X.; Nie, Y.; Shen, P.; Tan, S. Low-Cost Dyes Based on Methylthiophene for High-Performance Dye-Sensitized Solar Cells. *Dyes Pigm.* **2010**, *87*, 181–187.

(15) Mishra, A.; Ma, C.-Q.; Segura, J. L.; Bäuerle, P. Functional Oligothiophene-Based Materials: Nanoarchitectures and Applications. *Handbook of Thiophene-Based Materials* **2009**, 1–155.

(16) Vijay Kumar, C.; Cabau, L.; Koukaras, E. N.; Siddiqui, S. A.; Sharma, G. D.; Palomares, E. Efficient Bulk Heterojunction Solar Cells Based on Solution Processed Small Molecules Based on The Same Benzo[1,2-b:4, 5-b']Thiophene Unit as Core Donor and Different Terminal Units. *Nanoscale* **2015**, *7*, 7692–7703.

(17) Kim, D.; Salman, S.; Coropceanu, V.; Salomon, E.; Padmaperuma, A. B.; Sapochak, L. S.; Kahn, A.; Brédas, J.-L. Phosphine Oxide Derivatives as Hosts for Blue Phosphors: A Joint Theoretical and Experimental Study of Their Electronic Structure. *Chem. Mater.* **2010**, *22*, 247–254.

(18) Sapochak, L. S.; Padmaperuma, A. B.; Vecchi, P. A.; Cai, X.; Burrows, P. E. Designing Organic Phosphine Oxide Host Materials Using Heteroatomic Building Blocks: Inductive Effects on Electroluminescence. *Proc. SPIE* **2007**, 665506.

(19) Ljubić, I.; Sabljic, A. CASSCF/CASPT2 and TD-DFT Study of Valence and Rydberg Electronic Transitions in Fluorene, Carbazole, Dibenzofuran, and Dibenzothiophene. *J. Phys. Chem. A* **2011**, *115*, 4840–4850.

(20) Wex, B.; Kaafarani, B. K.; Danilov, E. O.; Neckers, D. C. Altering the Emission Behavior with the Turn of a Thiophene Ring: The Photophysics of Condensed Ring Systems of Alternating Benzenes and Thiophenes. *J. Phys. Chem. A* **2006**, *110*, 13754–13758.

(21) Qiao, Y.; Wei, Z.; Risko, C.; Li, H.; Brédas, J.-L.; Xu, W.; Zhu, D. Synthesis, Experimental and Theoretical Characterization, and Field-Effect Transistor Properties of a New Class of Dibenzothiophene Derivatives: From Linear to Cyclic Architectures. *J. Mater. Chem.* **2012**, *22*, 1313–1325.

(22) Kim, J.; Cho, N.; Min Ko, H.; Kim, C.; Kwan Lee, J.; Ko, J. Push-Pull Organic Semiconductors Comprising of Bis-Dimethylfluorenyl Amino Benzo[b]Thiophene Donor and Various Acceptors for Solution Processed Small Molecule Organic Solar Cells. *Sol. Energy Mater. Sol. Cells* **2012**, *102*, 159–166.

(23) Guarnaccio, A.; D'Auria, M.; Racioppi, R.; Mattioli, G.; Bonapasta, A. A.; De Bonis, A.; Teghil, R.; Prince, K. C.; Acres, R. G.; Santagata, A. Thiophene-Based Oligomers Interacting with Silver Surfaces and the Role of a Condensed Benzene Ring. *J. Phys. Chem. C* **2016**, *120*, 252–264.

(24) Giertz, A.; Bässler, M.; Björneholm, O.; Wang, H.; Feifel, R.; Miron, C.; Karlsson, L.; Svensson, S.; et al. High Resolution C1s and S2p Photoelectron Spectra of Thiophene. *J. Chem. Phys.* **2002**, *117*, 7587–7592.

- (25) Hitchcock, A. P.; et al. Inner Shell Excitation of Thiophene and Thiolane: Gas, Solid, and Monolayer States. *J. Chem. Phys.* **1986**, *85*, 4835–4848.
- (26) Grazioli, C.; Baseggio, O.; Stener, M.; Fronzoni, G.; de Simone, M.; Coreno, M.; Guarnaccio, A.; Santagata, A.; D'Auria, M. Study of The Electronic Structure of Short Chain Oligothiophenes. *J. Chem. Phys.* **2017**, *146*, 054303.
- (27) Baseggio, O.; Toffoli, D.; Stener, M.; Fronzoni, G.; de Simone, M.; Grazioli, C.; Coreno, M.; Guarnaccio, A.; Santagata, A.; D'Auria, M. S2p Core Level Spectroscopy of Short Chain Oligothiophenes. *J. Chem. Phys.* **2017**, *147*, 244301.
- (28) Norman, P.; Dreuw, A. Simulating X-ray Spectroscopies and Calculating Core-Excited States of Molecules. *Chem. Rev.* **2018**, *118*, 7208–7248.
- (29) Stöhr, J. *NEXAFS Spectroscopy*; Springer Verlag: Berlin, Heidelberg, Germany, and New York, 1992.
- (30) Blyth, R.; Delaunay, R.; Zitnik, M.; Krempasky, J.; Krempaska, R.; Slezak, J.; Prince, K. C.; Richter, R.; Vondracek, M.; Camilloni, R.; et al. The High Resolution Gas Phase Photoemission Beamline, Elettra. *J. Electron Spectrosc. Relat. Phenom.* **1999**, *101-103*, 959–964.
- (31) Mårtensson, N.; Baltzer, P.; Brühwiler, P. A.; Forsell, J.-O.; Nilsson, A.; Stenborg, A.; Wannberg, B. A Very High Resolution Electron Spectrometer. *J. Electron Spectrosc. Relat. Phenom.* **1994**, *70*, 117–128.
- (32) Myrseth, V.; Bozek, J.; Kukk, E.; Sæthre, L.; Thomas, T. Adiabatic and Vertical Carbon 1s Ionization Energies in Representative Small Molecules. *J. Electron Spectrosc. Relat. Phenom.* **2002**, *122*, 57–63.
- (33) Pettersson, L.; Nordgren, J.; Selander, L.; Nordling, C.; Siegbahn, K.; Ågren, H. Core-Electron Binding Energies in The Soft X-Ray Range Obtained in X-Ray Emission. *J. Electron Spectrosc. Relat. Phenom.* **1982**, *27*, 29–37.
- (34) True, J. E.; Thomas, T. D.; Winter, R. W.; Gard, G. L. Electronegativities from Core-Ionization Energies: Electronegativities of SF<sub>5</sub> and CF<sub>3</sub>. *Inorg. Chem.* **2003**, *42*, 4437–4441.
- (35) XPST tool download: <https://www.wavemetrics.com/project/XPSTools>.
- (36) Tronc, M.; King, G. C.; Read, F. H. Nitrogen K-Shell Excitation in N<sub>2</sub>, NO and N<sub>2</sub>O by High-Resolution Electron Energy-Loss Spectroscopy. *J. Phys. B: At. Mol. Phys.* **1980**, *13*, 999–1008.
- (37) 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.
- (38) Eustatiu, I. G.; Francis, J. T.; Tyliszczak, T.; Turci, C. C.; Kilcoyne, A. L. D.; Hitchcock, A. P. Generalized Oscillator Strengths for Inner-Shell Excitation of SF<sub>6</sub> Recorded with a High-Performance Electron Energy Loss Spectrometer. *Chem. Phys.* **2000**, *257*, 235–252.
- (39) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.
- (40) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: a Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- (41) Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. Towards an Order-N DFT Method. *Theor. Chem. Acc.* **1998**, *99*, 391–403.
- (42) 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.
- (43) Bak, B.; Christensen, D.; Hansen-Nygaard, L.; Rastrup-Andersen, J. The Structure of Thiophene. *J. Mol. Spectrosc.* **1961**, *7*, 58–63.
- (44) Yamazaki, D.; Nishinaga, T.; Komatsu, K. Radical Cation of Dibenzothiophene Fully Annulated with Bicyclo[2.2.2]octene Units: X-ray Crystal Structure and Electronic Properties. *Org. Lett.* **2004**, *6*, 4179–4182.
- (45) 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.
- (46) Slater, J. C. Statistical Exchange-Correlation in the Self-Consistent Field. *Adv. Quantum Chem.* **1972**, *6*, 1–92.
- (47) Triguero, L.; Pettersson, 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.
- (48) Mijovilovich, A.; Pettersson, L. G. M.; Mangold, S.; Janousch, M.; Susini, J.; Salome, M.; de Groot, F. M. F.; Weckhuysen, B. M. The Interpretation of Sulfur K-Edge XANES Spectra: A Case Study on Thiophenic and Aliphatic Sulfur Compounds. *J. Phys. Chem. A* **2009**, *113*, 2750–2756.
- (49) Becke, A. D. J. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (50) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of The Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.
- (51) Fronzoni, G.; De Francesco, R.; Stener, M.; Decleva, P. Spin-Orbit Relativistic Calculations of The Core Excitation Spectra of SO<sub>2</sub>. *J. Chem. Phys.* **2007**, *126*, 134308.
- (52) Fronzoni, G.; Stener, M.; Decleva, P.; Wang, F.; Ziegler, T.; van Lenthe, E.; Baerends, E. J. Spin-Orbit Relativistic Time Dependent Density Functional Theory Calculations for The Description of Core Electron Excitations: TiCl<sub>4</sub> Case Study. *Chem. Phys. Lett.* **2005**, *416*, 56–63.
- (53) Wang, F.; Ziegler, T.; van Lenthe, E.; van Gisbergen, S.; Baerends, E. J. The Calculation of Excitation Energies Based on The Relativistic Two-Component Zeroth-Order Regular Approximation and Time-Dependent Density-Functional with Full Use of Symmetry. *J. Chem. Phys.* **2005**, *122*, 204103.
- (54) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.
- (55) Stener, M.; Fronzoni, G.; de Simone, M. Time Dependent Density Functional Theory of Core Electrons Excitations. *Chem. Phys. Lett.* **2003**, *373*, 115–123.
- (56) Casida, M. E. Time-Dependent Density Functional Response Theory for Molecules. *Recent Advances in Density Functional Methods* **1995**, *1*, 155.
- (57) 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.
- (58) Cederbaum, L. S.; Domcke, W.; Schirmer, J. Many-Body Theory of Core Holes. *Phys. Rev. A: At., Mol., Opt. Phys.* **1980**, *22*, 206–222.
- (59) Barth, A.; Cederbaum, L. S. Many-Body Theory of Core-Valence Excitations. *Phys. Rev. A: At., Mol., Opt. Phys.* **1981**, *23*, 1038–1061.
- (60) Decleva, 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.
- (61) van Lenthe, E.; Baerends, E. J. Optimized Slater-Type Basis Sets for The Elements 1–118. *J. Comput. Chem.* **2003**, *24*, 1142–1156.
- (62) Fouda, A. E. A.; Besley, N. A. Assessment of Basis Sets for Density Functional Theory-Based Calculations of Core-Electron Spectroscopies. *Theor. Chem. Acc.* **2018**, *137*, 1–11.
- (63) van Leeuwen, R.; Baerends, E. J. Exchange-Correlation Potential with Correct Asymptotic Behaviour. *Phys. Rev. A: At., Mol., Opt. Phys.* **1994**, *49*, 2421–2431.
- (64) Fronzoni, G.; Baseggio, O.; Stener, M.; Hua, W.; Tian, G.; Luo, Y.; Apicella, B.; Alfè, M.; de Simone, M.; Kivimaki, A.; et al. Vibrationally Resolved High-Resolution NEXAFS and XPS Spectra of Phenanthrene and Coronene. *J. Chem. Phys.* **2014**, *141*, 044313.
- (65) Zhang, T.; Brumboiu, I. E.; Grazioli, C.; Guarnaccio, A.; Coreno, M.; de Simone, M.; Santagata, A.; Rensmo, H.; Brena, B.; Lanzilotto, V.; et al. Lone-Pair Delocalization Effects within Electron

Donor Molecules: The Case of Triphenylamine and Its Thiophene-Analog. *J. Phys. Chem. C* **2018**, *122*, 17706–17717.

(66) Lin, L.-Y.; Chen, Y.-H.; Huang, Z.-Y.; Lin, H.-W.; Chou, S.-H.; Lin, F.; Chen, C.-W.; Liu, Y.-H.; Wong, K.-T. A Low-Energy-Gap Organic Dye for High-Performance Small-Molecule Organic Solar Cells. *J. Am. Chem. Soc.* **2011**, *133*, 15822–15825.

**SUPPORTING INFORMATION  
FOR PUBLICATION**

**Electronic Structure Characterization of a Thiophene Benzo-annulated Series of  
Common Building Blocks for Donor and Acceptor Compounds Studied by Gas Phase  
Photoelectron and Photoabsorption Synchrotron Spectroscopies**

*D. Toffoli*<sup>1</sup>, *A. Guarnaccio*<sup>2\*</sup>, *C. Grazioli*<sup>2</sup>, *T. Zhang*<sup>3</sup>, *F. Johansson*<sup>3</sup>, *M. de Simone*<sup>4</sup>, *M. Coreno*<sup>2</sup>, *A. Santagata*<sup>2</sup>, *M. D'Auria*<sup>5</sup> and *C. Puglia*<sup>3</sup>

*E. Bernes*<sup>1</sup>, *M. Stener*<sup>1</sup>, *G. Fronzoni*<sup>1\*</sup>

<sup>1</sup> Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy

<sup>2</sup> ISM-CNR, Tito Scalo (PZ) and Trieste LD2 Unit, Italy

<sup>3</sup> Department of Physics and Astronomy, Uppsala University, Box 516, SE-751 20 Uppsala, Sweden.

<sup>4</sup> IOM-CNR, Laboratorio TASC, Sincrotrone Trieste, I-34149 Trieste, Basovizza, Italy

<sup>5</sup> Science Department, University of Basilicata, 85100 Potenza, Italy

\*email: [ambra.guarnaccio@pz.ism.cnr.it](mailto:ambra.guarnaccio@pz.ism.cnr.it)

\*email: [fronzoni@units.it](mailto:fronzoni@units.it)

**TABLE S1** - Comparison between theoretical and experimental geometrical parameters for T. Bond lengths are expressed in Å, angles in degrees.

<b>Bond lengths</b>	Calculated	Experimental
S-C1	1.7339	1.7140
C1-C2	1.3645	1.3700
C2-C3	1.4274	1.4230
<b>Angles</b>	Calculated	Experimental
C-S-C	91.369	92.100
S-C1-C2	111.601	111.280
C1-C2-C3	112.715	112.270

**TABLE S2** - Comparison between theoretical and experimental geometrical parameters for DBT. Bond lengths are expressed in Å, angles in degrees.

<b>Bond lengths</b>	Calculated	Experimental
S-C1	1.7357	1.7490
C1-C4	1.3867	1.3920
C4-C6	1.3815	/
C6-C5	1.3947	/
C5-C3	1.3800	/
C3-C2	1.3911	/
C2-C1	1.4041	1.4130
<b>Angles</b>	Calculated	Experimental
C-S-C	91.450	91.790
S-C1-C4	126.18	126.17
S-C1-C2	112.29	112.15
C1-C4-C6	118.48	/
C4-C6-C5	120.76	/
C6-C5-C3	119.80	/
C5-C3-C2	120.48	/
C3-C2-C1	118.86	/
C2-C1-C4	121.59	/



**TABLE S3** - Comparison between theoretical  $\Delta$ KS C1s IPs computed at the DFT level by using the PW86xPerdew and the hybrid B3LYP xc potentials for all non-equivalent Ci atoms in BBT and DBT. All values are expressed in eV.

Site	BBT		DBT	
	PW86xPerdew	B3LYP	PW86xPerdew	B3LYP
<b>C1</b>	290.78	290.36	290.68	290.27
<b>C2</b>	290.49	290.09	290.33	289.93
<b>C3</b>	290.21	289.95	290.15	289.78
<b>C4</b>	290.33	290.00	290.24	289.88
<b>C5</b>	290.37	290.08	290.07	289.73
<b>C6</b>	290.18	289.82	290.10	289.76
<b>C7</b>	290.17	289.81	-	-
<b>C8</b>	290.67	290.28	-	-

# PPT Isolated Molecule and Its Building Block Moieties Studied by C 1s and O 1s Gas Phase X-ray Photoelectron and Photoabsorption Spectroscopies

Ambra Guarnaccio,\* Teng Zhang, Cesare Grazioli, Fredrik O. L. Johansson, Marcello Coreno, Monica de Simone, Giovanna Fronzoni,\* Daniele Toffoli, Elisa Bernes, and Carla Puglia

Cite This: *J. Phys. Chem. C* 2020, 124, 9774–9786

Read Online

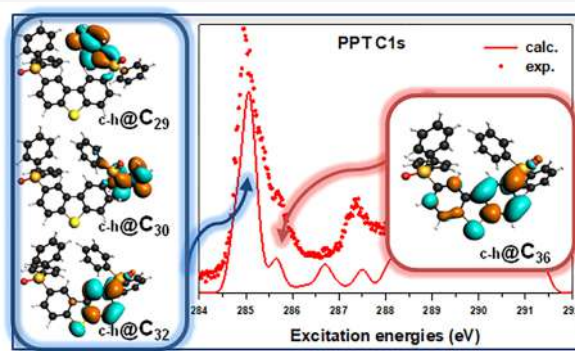
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** The present study is focused on the comprehensive gas phase electronic structure characterization of 2,8-bis-(diphenylphosphoryl)-dibenzo[b,d]thiophene (PPT), a promising ambipolar phosphorescent host material recently introduced in organic light-emitting diodes (OLEDs). This molecular system can be considered ideally formed by two diphenylphosphine oxide (dPPO) moieties functionalizing the small dibenzothiophene (DBT) core. PPT is characterized by high triplet energy and is known as good vacuum sublimable electron transporting material for blue OLEDs. The triphenyl phosphine oxide (TPPO) molecule has been chosen as the model compound of the dPPO groups in PPT. A combined experimental and theoretical study by density functional theory of the gas phase electronic structure of TPPO and PPT has been performed through X-ray photoelectron spectroscopy and near-edge X-ray absorption fine structure spectroscopy measured at the carbon and oxygen 1s regions. The study represents a detailed characterization of the impact of the single building blocks on the electronic structure of the whole PPT molecule. Moreover, it confirms that the phosphine oxide groups act as breaking points of the  $\pi$ -conjugation between the DBT core of PPT and the outer groups, leaving the electronic structures of the compound practically matching those of the central DBT moiety.



## INTRODUCTION

In organic electronics, homojunction devices represent from a commercial point of view an attractive alternative to the more traditional heterojunction systems due to their simplicity and ease of processing. In contrast to the heterojunction structures sketched in Figure 1a, where different layers of materials with specific properties allow the charge transport and recombination, the homojunction organic devices (as, for example, organic light-emitting diodes (OLEDs)), represented in Figure 1b, are based on few organic materials with the multiple roles of hole/electron transport and light emission. This requires usually an ambipolar molecular film, which allows high and balanced mobility of both holes and electrons. To accomplish this, the ambipolar organic material is made of different functional groups inserted inside the same molecular system. In general, it is composed of three basic building blocks: a hole transporting molecule (donor), an electron transporting system (acceptor), and a polycyclic aromatic moiety used as spacer, as shown in Figure 1c. The spacer is important to fine-tune the properties since the increased size promotes a redshift of the emission spectrum and also improves the device efficiency as a whole.<sup>1</sup>

This study reports the characterization of a promising ambipolar material, PPT (2,8-bis-(diphenyl-phosphoryl)-dibenzo[b,d]thiophene,  $C_{36}H_{26}O_2P_2S$ ), shown in Figure 2a. PPT has already been implemented as the host of blue light-emitting materials, such as FIrPic (bis[2-(4,6-difluorophenyl)pyridinato-C2,N](picolinato)iridium), with sky-blue emission, high emission efficiency, and suitable energy levels as a phosphorescent dopant material.<sup>2,3</sup> Moreover, PPT has been also used in combination with electron-donating materials to enhance the electroluminescence efficiency.<sup>4–7</sup>

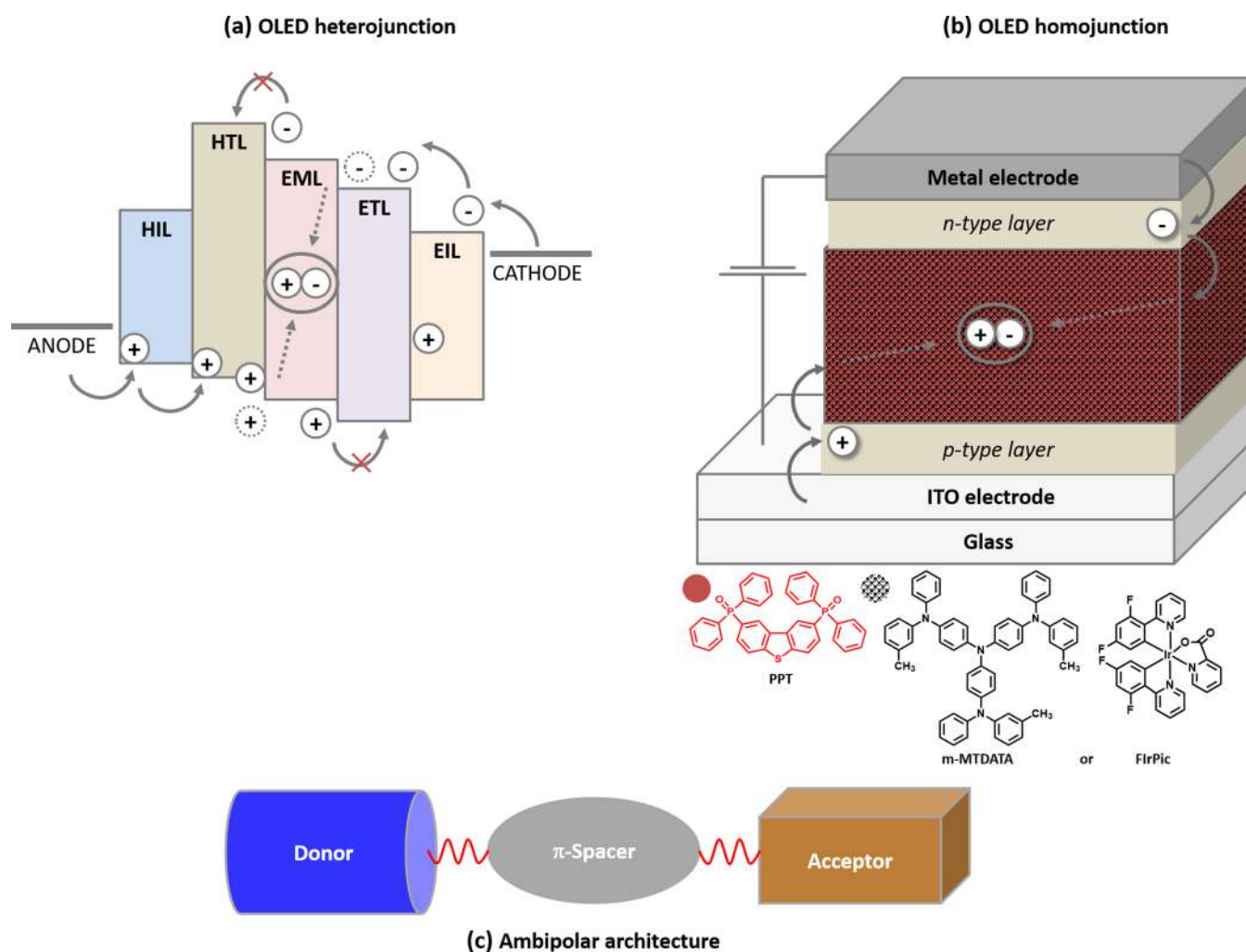
PPT is characterized by a wide bandgap with high triplet energy level. It is composed by an electron-rich dibenzothiophene (DBT) core and two electron-deficient diphenylphosphoryl (dPPO) side arms, ensuring its good electron- and hole-transporting property, thus maintaining charge balance in the

Received: February 28, 2020

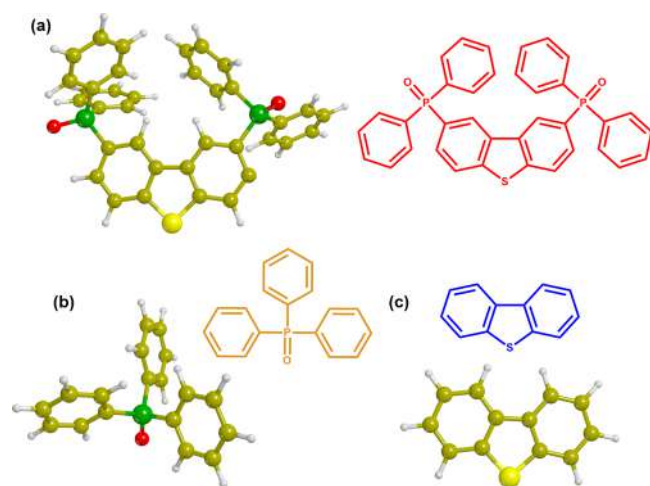
Revised: April 8, 2020

Published: April 10, 2020





**Figure 1.** (a) Sketch of a heterojunction OLED device architecture. The flow of the negative charge carriers (electrons) from the cathode and the positive charge carriers (holes) from the anode have been indicated by the arrows until the final recombination in the EML. Each specific layer has been labeled as hole injection layer (HIL), hole transport layer (HTL), electron injection layer (EIL) and electron transport layer (ETL). (b) Sketch of a homojunction OLED device architecture. (c) General structure of an ambipolar molecule used as host material.



**Figure 2.** Chemical and optimized molecular structures of (a) 2,8-bis(diphenyl-phosphoryl)-dibenzo[b,d]thiophene (PPT), (b) triphenylphosphine oxide (TPPO), and (c) dibenzothiophene (DBT) molecules. The colored balls considered in the DFT calculations indicate the different atoms: carbon (beige), phosphor (green), sulfur (yellow), and oxygen (red).

emissive layer (EML) of phosphorescent organic light-emitting diodes (PhOLEDs).<sup>8,9</sup> It has been also shown that the performance of devices implementing ambipolar materials is especially improved when, among other analogues cores (i.e., carbazole and dibenzofuran), the  $\pi$ -conjugated DBT is used.<sup>10</sup> This polycyclic aromatic building block in the PPT core promotes the hole transport and the lowering of the HOMO and LUMO levels facilitating the electron injection process and introducing a hole-blocking function in the OLED.

The phosphine oxide (PO), belonging to a known electron-withdrawing class of materials,<sup>11–14</sup> represents instead the electron-transport moiety in PPT. The good electron mobility of PO-based materials allows them to reach a good charge balance in the recombination zone. Furthermore, phosphine oxides have high triplet energy that helps to prevent the back-energy transfer from the emitter to the host. The introduction of PO group(s) is also beneficial for the thermal stability of the resulting molecular system (in our case PPT).<sup>15</sup> Indeed, it has been shown that building blocks such as phosphine oxides (small, volatile, and with high triplet state energy) can be used to make larger molecules more amenable to sublimation despite their significant size.<sup>15</sup> It has been also suggested<sup>12</sup> that the PO groups in PPT act as breaking points of  $\pi$ -conjugation

between the core and the outer side arms of the molecular system. This would mean that the PO moieties have practically no impact on the PPT electronic structure which would then mostly resemble that of the DBT core.

Among the cited studies, the work by Kim et al. is mainly focused on a joint theoretical and experimental investigation of the electronic structure of a series of bis(diphenylphosphine oxide) derivatives containing as the central aromatic core, among others, also the DBT moiety. With respect to the isolated DBT molecule, the addition of the diphenylphosphine oxide arms was revealed to only slightly reduce the energy of the lowest triplet state. Moreover, in the cited study the ultraviolet photoelectron spectroscopy (UPS) and inverse photoemission spectroscopy (IPES) characterizations, using conventional He lamps, indicated that different diphenylphosphine oxide functionalities impact significantly only on the ionization potential and the electron affinity of the molecular system.

Our study represents the first combined theoretical and experimental investigations which accurately map the electronic structure, looking at the occupied and virtual states by means of synchrotron photoelectron and photoabsorption spectroscopies of free PPT and triphenylphosphine oxide (TPPO). In this study, TPPO is considered as the building block of the peripheral dPPO arms in PPT. These results were also compared with previous characterizations of dibenzothio-phenene (DBT),<sup>16</sup> which instead represents the PPT core. All these molecules are sketched in Figure 2a–c.

TPPO and PPT molecules have been probed in the gas phase by means of X-ray photoelectron spectroscopy (XPS) and near-edge X-ray absorption fine structure (NEXAFS) spectroscopy and the experimental results have been supported by DFT calculations; these are all well-known powerful tools to thoroughly investigate the electronic structure of free molecules due to the localized character of the core–hole. The aim is to investigate and identify the impact of the single building blocks on the electronic structure of PPT and to spectroscopically confirm the predicted conjugation breaking between the DBT core and the dPPO arms. The study of the electronic structure of technologically relevant conjugated complexes compared to their building blocks (in our case, PPT compared to TPPO and DBT) has already turned out to be a very successful strategy<sup>16–18</sup> and provides new knowledge for the design of novel molecular structures to be implemented in high-performance devices.

## ■ EXPERIMENTAL METHODS

Commercially available Sigma-Aldrich triphenylphosphine oxide (TPPO, 98%) and 2,8-bis(diphenyl-phosphoryl)-dibenzo[b,d]thiophene (PPT, >99%) from Lumtec were used. PPT and TPPO are crystalline solid powders, and they were sublimated in vacuum using a custom built resistively heated furnace. The best vaporization temperature for PPT was found to be 251 °C after a temperature gradient sublimation treatment to maximum 200 °C for 24 h. TPPO was sublimated at 128 °C after a gradient temperature sublimation up to 100 °C kept for 20 h. The experimental details of the characterization of DBT molecule are reported in the paper by Toffoli et al.<sup>16</sup>

XPS measurements on gas phase samples were performed at the Gas Phase PhotoEmission beamline of the Elettra Synchrotron in Trieste<sup>19</sup> using a Scienta SES-200 electron

analyzer<sup>20</sup> mounted at the magic angle with respect to the electric vector of the linearly polarized incident light.

C 1s XP spectra of PPT and TPPO molecules were recorded at a photon energy of 392 eV with an overall energy resolution of about 200 meV. The binding energy (BE) scale was calibrated with respect to the C 1s binding energy (BE) of CO<sub>2</sub> (297.6 eV).<sup>21</sup>

A photon energy of 628 eV was used for the O 1s XP spectra of PPT and TPPO with an overall energy resolution of about 400 meV. The spectra were calibrated with respect to the O 1s binding energy of CO<sub>2</sub> (O 1s 541.2 eV).<sup>22</sup>

The C 1s and O 1s XPS results were fitted using IGOR PRO with the package XPST- X-ray Photoelectron Spectroscopy Tools by Dr. Martin Schmid. The fitting of the experimental C 1s and O 1s XP spectra, shown in the following sections, was performed considering a linear background for both PPT and TPPO molecules and a pseudo-Voigt profile<sup>23</sup> with a Gaussian–Lorentzian (G–L) ratio equal to 0.150 eV for TPPO and a pure Gaussian profile (G–L = 0) for PPT. Moreover, an additional asymmetry coefficient of 0.304 and a width of 0.400 eV were applied to consider the line profile of benzene-like carbons in the case of the smaller TPPO molecule. The same width but an asymmetry coefficient of 0.001 was used for the PPT molecule with a more sterically hindered core and then characterized by a less significant vibrational progression.

For fitting the C 1s XP spectra, at each inequivalent carbon atom was assigned a fit component with a relative intensity according to the molecular stoichiometry and a binding energy (BE) taken from the calculated ionization potentials (IPs). After the fitting, a rigid shift was performed to better align the IP values with the resulting fitted components. To prove the validity of the theoretical IP values, we considered the components used for fitting the experimental spectra as a convolution of the calculated IPs by pseudo-Voigt curves. However, the binding energies of the components of the final fit are those that better describe the experimental data. Their only minor divergence from the calculated IPs still confirms the good agreement between the experimental and theoretical results.

The fittings of the experimental O 1s spectra were performed considering a linear background for both PPT and TPPO systems. In both cases, one pseudo-Voigt profile<sup>23</sup> was used with a G–L ratio equal to 0.300 eV for TPPO and PPT. The asymmetry coefficients of 0.001 and 0.149 and widths of 0.800 and 0.700 eV were used for PPT and TPPO, respectively.

However, the fits of the XPS experimental data are used in this work just to give a qualitative confirmation of the IPs obtained by calculations. The fits are not intended for a conclusive description of the experimental data due to the many different chemical inequivalent sites/atoms contributing to the spectral intensity.

NEXAFS spectra at the C 1s were measured recording the total electron yield collected by a channeltron. The O 1s NEXAFS spectrum was measured following the Auger yield using the SES200 electron analyzer in a fixed kinetic energy window mode. All the NEXAFS spectra were normalized by the transmitted photon flux measured by a calibrated Si photodiode.

The energy scales of the C 1s NEXAFS and O 1s NEXAFS spectra were calibrated by taking simultaneous spectra of the samples and of CO<sub>2</sub> with the characteristic transition at 290.77

eV (C 1s  $\rightarrow \pi^*$ , CO<sub>2</sub>)<sup>24,25</sup> and at 535.30 eV (O 1s  $\rightarrow \pi^*$ , CO<sub>2</sub>), respectively. The photon energy resolution of the C 1s NEXAFS spectra was around 110 meV for PPT and TPPO, while it was 65 meV for DBT molecules. Moreover, the photon energy resolution of the O 1s NEXAFS spectra was around 120 meV.

## COMPUTATIONAL DETAILS

**Geometry Optimizations.** The equilibrium geometries of TPPO and PPT have been optimized at the density functional theory (DFT)<sup>26</sup> level, within the local-density approximation (LDA) by using the Vosko, Wilk, Nusair (VWN) exchange and correlation (xc) functional<sup>27</sup> and the triple  $\zeta$  polarized (TZP) basis set of Slater type orbitals (STOs) extracted from the ADF (Amsterdam Density Functional) database.<sup>28,29</sup> For both molecules, the geometry optimization computations have been performed without imposing any symmetry constraint. Crystallographic studies have revealed that TPPO as a whole has no symmetry because the rotations of the phenyl rings about the P—C bonds are not identical.<sup>30</sup> In addition, due to the partial multiple bond character of the P=O bond, the O=P—C angles result are greater than the tetrahedral angle, whereas the C—P—C angles are smaller than it.<sup>31</sup> Analogous considerations can be made in the case of PPT since, due to steric hindrance, the two diphenylphosphine oxide arms are not specular with respect to the half molecular skeleton and point out toward different directions. In particular, in the case of the TPPO molecule, two different conformations have been considered: more specifically, a structure from X-ray diffraction (XRD) data,<sup>32</sup> and one predicted by using the MP2/cc-pVTZ level of theory in the gas phase.<sup>33</sup> The investigation of the two different conformations has been supported by a work by Al-Farhan.<sup>31</sup> From a comparison between the geometric parameters (i.e., bond lengths, bond angles, and torsional angles) of the two conformers (see Table S1 of the Supporting Information, SI), one can notice that the two sets of data present similar values in particular as concerns both bond lengths and bond angles. Significant variations are instead found in torsional angles, since the corresponding deviations are of the order of about 40° (see Table S1); this is supported by Figure S1 from which one can notice that the phenyl rings in the two conformers are differently twisted with respect to the P=O bond axis. Therefore, in the calculations of the NEXAFS spectra, both conformers have been taken into account. Here, the presented calculated data refer to conformer *a* (see Figure S1a). As NEXAFS spectra of the two conformers are quantitatively very similar (see Figure S2), we chose to carry out the assignment only on the structure derived from XRD data.

**XPS and NEXAFS Calculations.** The calculations of the XPS and NEXAFS spectra have been performed employing the DFT method and the generalized gradient approximation (GGA) for the xc functional (PW86xPerdew)<sup>34</sup> as implemented in the ADF program.<sup>28</sup> The simulation of the XP spectra has been performed by computing the  $\Delta$ KS ( $\Delta$ SCF Kohn–Sham) C1 s<sup>-1</sup> and O1 s<sup>-1</sup> IPs. The energy of the 1 s<sup>-1</sup> ionic state has been obtained through a KS unrestricted calculation. All the calculated spectral profiles have been convoluted by using Gaussian functions of appropriate full width at half maximum (fwhm) value (in C 1s NEXAFS spectra, equal to 0.3 eV and 0.4 eV for TPPO and PPT, respectively, while in O 1s NEXAFS spectra, equal to 0.3 and 0.5 eV for TPPO and PPT, respectively) as giving the best fit

to the experimental data. Moreover, the theoretical line-shapes in C 1s NEXAFS spectra have been shifted by 0.63 and of 0.58 eV to lower energies (for TPPO and PPT, respectively), while in O 1s NEXAFS spectra, a rigid shift of 0.3 and 0.6 eV to lower energies has been applied to get an easier comparison with the experimental data. To support the interpretation of the IP trend, we performed a Bader's population analysis by employing the PW86x Perdew xc potential.<sup>34</sup> Bader atomic properties have been calculated by using a grid-based method reported by Rodriguez et al.<sup>35,36</sup> and implemented in the ADF program.<sup>28</sup>

The C and O 1s NEXAFS spectra have been obtained by performing a separate computation of the excitation spectrum for each nonequivalent C/O site, then the total spectrum has been obtained by summing up the partial contributions weighted by the number of equivalent C/O atoms. Each spectrum has been computed at DFT level employing the Transition Potential (TP) approach<sup>37,38</sup> in which half an electron is removed from the excited core orbital, leaving all the virtual orbitals unoccupied and relaxing all the orbitals until self-consistency is obtained. Usually this approach adequately describes most of the relaxation effects upon formation of the core–hole and provides a single set of orthogonal orbitals from which transition dipole moments can be obtained.<sup>39</sup> The TP scheme is well-known for its ability to simulate K-shell NEXAFS spectra of light atoms.<sup>40</sup> An even tempered quadruple- $\zeta$  with three polarization and three diffuse functions (designed as ET-QZ3P-3DIFFUSE set in the ADF database) has been employed for the core-excited C and O atoms, in order to accurately describe the higher energy excitations toward diffuse Rydberg states which contribute to the near-edge spectral features.<sup>41</sup> It is well-known that the use of a good basis set enlarged with diffuse functions on the excited atom allows for an improved representation of the relaxation effects of the inner orbitals.<sup>41</sup>

Furthermore, the frozen core (FC) approximation has been used for core orbitals of the nonexcited atoms; it consists of treating explicitly only outer level electrons, while the innermost (core) atomic shells are kept frozen. This technique ensures the localization of the half core–hole as well as a further reduction of the computational effort without losing in accuracy. In particular, a FC TZP.1s basis set has been employed for the C and O atoms, while a FC TZP.2p basis set has been adopted for the P and S atoms.

Within the TP scheme, excitation energies are obtained as the differences between the eigenvalues of the virtual orbital and that of the 1s core orbital calculated with the TP configuration

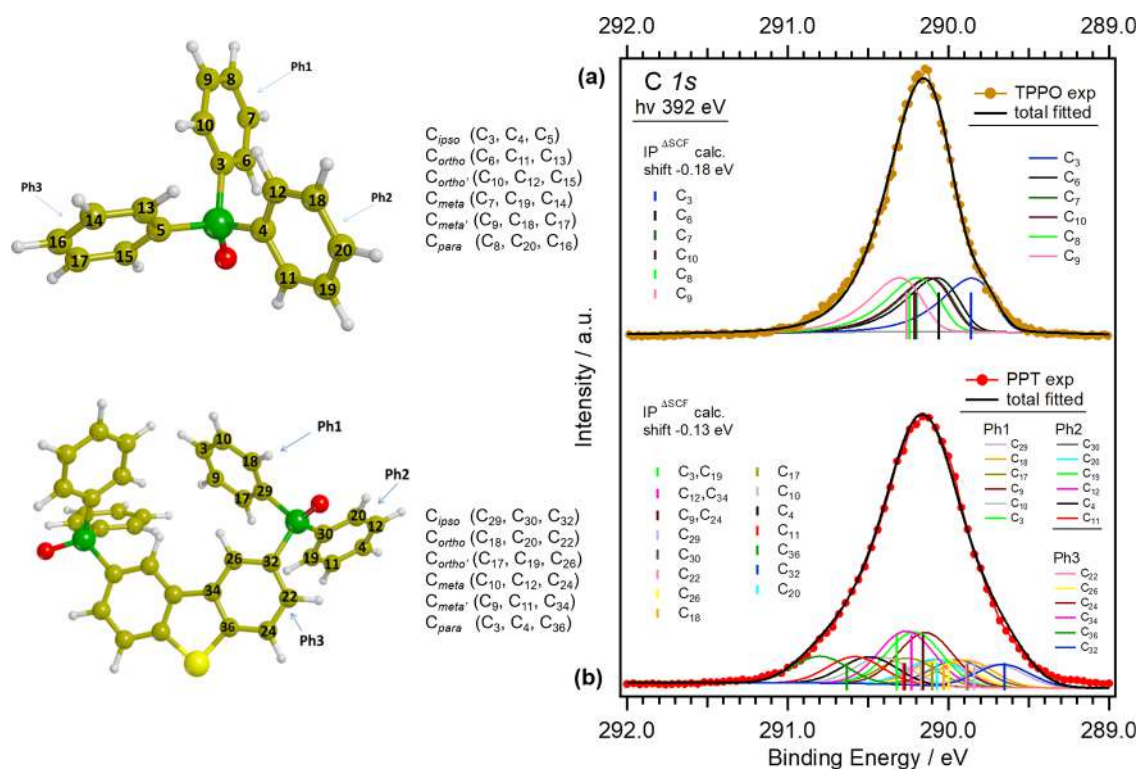
$$\Delta E_{i \rightarrow f} = \epsilon_f^{\text{TP}} - \epsilon_i^{\text{TP}} \quad (1)$$

while transition intensities are expressed in terms of oscillator strengths,  $f_{i \rightarrow f}$ . For samples in gas phase, as those investigated in this study,  $f_{i \rightarrow f}$  reads

$$f_{i \rightarrow f} = \frac{2}{3} n_i \Delta E_{i \rightarrow f} |\langle \varphi_f^{\text{TP}} | \boldsymbol{\mu} | \varphi_i^{\text{TP}} \rangle|^2 \quad (2)$$

involving dipole matrix elements between initial and final TP MOs, where  $n_i$  denotes the occupation number of the core orbital in the ground state.

The ionization potential (IP) is defined as the negative of the TP eigenvalue related to the initial core orbital,  $\text{IP} = -\epsilon_f^{\text{TP}}$ . Since the TP approach leads to a less attractive potential and



**Figure 3.** C 1s photoelectron spectra of TPPO (a) and PPT (b). Dotted curves represent the experimental data, while solid black lines represent the total fit results. Each fitted  $C_i$  component is represented by the colored solid curves, while the theoretical IPs are reported as vertical bars. Colors follow the attribution of the fitted components to the theoretical IPs reported in Table 1. On the left we report the three-dimensional molecular structures with the classification of the carbon atoms as discussed in the text.

**Table 1. Comparison between Theoretical IPs and Experimental BEs of Nonequivalent C Atoms of TPPO and PPT Obtained by a Fit of the C 1s PE Spectrum<sup>a</sup>**

	IP - $\Delta K S^{b,c}$						BE (fit) TPPO	BE (fit) PPT
	Ph1		Ph2		Ph3			
	TPPO	PPT	TPPO	PPT	TPPO	PPT		
$C_{ippo}$	(C <sub>3</sub> ) 289.86	(C <sub>29</sub> ) 289.84	(C <sub>4</sub> ) 289.86	(C <sub>30</sub> ) 289.88	(C <sub>5</sub> ) 289.85	(C <sub>32</sub> ) 289.66	289.85	289.70 289.89 289.66
$C_{ortho}$	(C <sub>6</sub> ) 290.06	(C <sub>18</sub> ) 290.03	(C <sub>11</sub> ) 290.06	(C <sub>20</sub> ) 290.07	(C <sub>13</sub> ) 290.06	(C <sub>22</sub> ) 289.89	290.07	290.02 290.05 289.94
$C_{ortho'}$	(C <sub>10</sub> ) 290.21	(C <sub>17</sub> ) 290.10	(C <sub>12</sub> ) 290.21	(C <sub>19</sub> ) 290.19	(C <sub>15</sub> ) 290.21	(C <sub>26</sub> ) 290.00	290.13	290.25 290.21 290.08
$C_{meta}$	(C <sub>7</sub> ) 290.20	(C <sub>10</sub> ) 290.17	(C <sub>19</sub> ) 290.21	(C <sub>12</sub> ) 290.23	(C <sub>14</sub> ) 290.21	(C <sub>24</sub> ) 290.16	290.12	290.41 290.14 290.26
$C_{meta'}$	(C <sub>9</sub> ) 290.26	(C <sub>9</sub> ) 290.16	(C <sub>18</sub> ) 290.26	(C <sub>11</sub> ) 290.28	(C <sub>17</sub> ) 290.26	(C <sub>34</sub> ) 290.23	290.31	290.26 290.49 290.14
$C_{para}$	(C <sub>8</sub> ) 290.24	(C <sub>3</sub> ) 290.19	(C <sub>20</sub> ) 290.25	(C <sub>4</sub> ) 290.27	(C <sub>16</sub> ) 290.24	(C <sub>36</sub> ) 290.63	290.19	290.21 290.58 290.80

<sup>a</sup>All values are expressed in eV. <sup>b</sup>The calculated IPs have been shifted by  $-0.18$  eV (TPPO). <sup>c</sup>The calculated IPs have been shifted by  $-0.13$  eV (PPT) to match the experimental peaks.

consequently the absolute transition energies are generally too large, the NEXAFS energies have been adjusted by shifting the TP excitation energies (eq 1) with respect to the  $\Delta IP$  value which corresponds to the energy difference  $\epsilon_{1s}^{TP} - \Delta KS(1s)$ .

The energy of the  $\Delta KS 1 s^{-1}$  ionic state corresponds to that previously calculated for the XP spectra.

In each NEXAFS spectrum, theoretical ionization thresholds are shown; these are mainly useful to separate the below-edge

region, where a discrete orbital description is absolutely adequate from the above-edge region where only qualitative information can be extracted, as a consequence of the discretization of the nonresonant continuum that is an artifact of the calculation.

## RESULTS AND DISCUSSION

**Experimental and Theoretical PES Synchrotron Analysis. C 1s Core Level.** Experimental photoelectron spectra of the C 1s core levels of TPPO and PPT are displayed in Figure 3a,b with dotted curves, together with the analysis of the experimental line-shapes fitted (solid black line) with the procedure described in Experimental Methods. The fit results are compared with the theoretical  $\Delta$ KS ionization potentials (IPs) for the nonequivalent carbon atoms reported in Figure 3 as vertical bars of different colors. Each fitted component (displayed as colored solid curves) are attributed to the theoretical IPs; the colors follow the attribution reported in Table 1. For a better agreement with the experimental fit components, all the calculated IPs have been shifted by  $-0.18$  eV (TPPO) and by  $-0.13$  eV (PPT). As it can be seen, the results of the experimental fit are quite well in agreement with the calculated IPs ( $\Delta$ KS).

The discussion of the results is complicated by the large number of different carbon atoms. We have therefore classified the carbon atoms ( $C_i$ ) of each phenyl group (Ph1, Ph2, and Ph3) on the basis of their positions, as reported in Figure 3. In particular, the carbon ipso, ortho, meta, and para positions have been defined with respect to the phosphorus atom. Furthermore, the ortho/ortho' and meta/meta' positions of the C atoms have been defined according to their spatial distance from the O atom. In particular,  $C_{\text{ortho}}$  and  $C_{\text{meta}}$  denote the C atoms closer to oxygen with respect to the  $C_{\text{ortho}'}$  and  $C_{\text{meta}'}$ , respectively. For the TPPO molecule, these C—O spatial distances are listed in Table S2 of the SI. The TPPO molecule has no symmetry, however the analysis of the geometrical parameters (see Table S1 of the SI) indicates that the C atoms occupying the same relative positions (ipso, ortho/ortho', meta/meta', and para) in the three phenyl rings are equivalent. We note that the positions ortho/ortho' on each phenyl ring are nonequivalent; the same applies for the positions meta/meta'. The same classification has been adopted for the PPT system where the ortho/ortho' and meta/meta' positions are still distinguished according to their spatial distance from the O atom of the dPPO moiety. Table 1 compares the BEs obtained by the fit of the experimental spectrum (BE (fit)) and the theoretical IPs (IP  $- \Delta$ KS) for the TPPO and PPT molecules.

**TPPO C 1s Core Level.** The calculated IP values of TPPO reported in Table 1 reflect the substantial equivalence of the three phenyl rings labeled as Ph1, Ph2, and Ph3 (see left panel of Figure 3a) therefore, only six distinct IPs contribute to the XP spectrum. Two of them, corresponding to the  $C_{\text{ortho}'}$  and  $C_{\text{meta}'}$  positions, are substantially degenerate (IP values are 290.21 or 290.20 eV in Table 1).

To assist in the interpretation of the observed trend, we report in Table S3 the SI results of a Bader charge analysis from which it can be clearly seen that a large negative charge is localized on the  $C_{\text{ipso}}$  while the charge on all other C atoms is very small in magnitude. This explains the lower calculated IP for the  $C_{\text{ipso}}$ . The lower IPs (of  $C_{\text{ortho}'}$ ,  $C_{\text{meta}'}$ ,  $C_{\text{meta}}$  and  $C_{\text{para}}$ ), instead, are clustered in a narrow energy range and from the population analysis alone it is difficult to rationalize their

values. The comparison between shifted theory and experiment fit values reveals a general good agreement with deviations within 10–90 meV since the IPs values are slightly overestimated by theory by about 180 meV (see Table 1 and Figure 3).

**PPT C 1s Core Level.** The XPS results of the PPT molecule are reported in Figure 3b and in Table 1. Looking at the geometrical parameters of the PPT (Table S2 of the SI), the three phenyl rings Ph1, Ph2, and Ph3 (left panel of Figure 3b) are equivalent to the corresponding rings on the left-hand side of the thiophene in the DBT moiety. We therefore calculated the IPs only for the  $C_i$  atoms of the Ph1, Ph2, and Ph3 phenyl rings. All these carbons are nonequivalent and have been classified on the basis of their position in the phenyl ring (see the inset of Figure 3b), as previously described.

To identify and explain the inequivalent carbon sites in PPT, we need to consider their different geometrical surrounding and the chemical bond with different heteroatoms, like S and P, both characterized by a higher electronegativity than carbon. The different geometrical surrounding leads us to ideally separate the phenyl groups that are free (Ph1 and Ph2) from those condensed with the thiophene ring (Ph3) as shown in Figure 3b. Moreover, because of the existence of two different kinds of heteroatoms in PPT, within the previously described phenyl systems we can further differentiate carbons depending on their proximity to the sulfur (Ph3) or to the phosphorus (Ph1, Ph2, and also Ph3) heteroatoms. In this context,  $C_{\text{para}}$  ( $C_{36}$ ) is the carbon directly bonded to the sulfur, and  $C_{\text{ipso}}$  are the carbons ( $C_{29}$ ,  $C_{30}$ , and  $C_{32}$ ) directly bonded to the phosphorus atom.

Some degeneracies are found among the calculated IP values of the 18 nonequivalent  $C_i$ 's reported in Table 1 for PPT, therefore only 15 distinct vertical IPs are represented by the bars reported in the plot of Figure 3b.

The IPs in Ph1 and Ph2 increase from  $C_{\text{ipso}}$  to  $C_{\text{para}}$  as already found in the TPPO molecule: they preserve their electronic environment inside the PPT molecule, and the presence of the DBT moiety influences their IPs only to a limited extent.

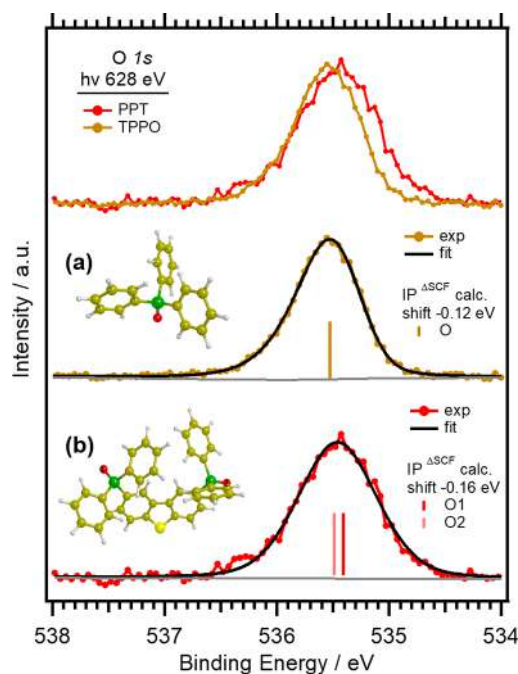
For the Ph3 ring, the counteracting inductive and resonance effects come into play in determining the trend of the calculated IPs of the  $C_i$  sites. The inductive effects due to the oxygen of the PO moiety would tend to increase the IP of the closest  $C_i$  atoms. On the other hand, resonance effects of the S  $p_z$  lone pairs in the central DBT core also play a role as shown in our previous publication<sup>16</sup> and as confirmed also by the calculated Bader charges shown in Table S3, tending to lower the IP values of the Ph3 annulated ring, stabilizing in particular the  $C_{32}$  (ipso) and the  $C_{24}$  (meta) sites.

Moreover, the IP of the  $C_{32}$  (ipso) is similar to the value found for the corresponding atom of free DBT,<sup>16</sup> indicating that the electronic structure of the PPT core is not perturbed by the nearby PO groups, confirming a previous study.<sup>15</sup>

As expected by considering the proximity of the higher electronegative S atom,  $C_{36}$  (para) results show the highest IP. Among the C atoms belonging to Ph3, after  $C_{36}$  (para) in decreasing IP order we found the  $C_{34}$  (meta'). The IPs of the remaining  $C_i$  of the annulated Ph3 are slightly different from the values of the corresponding carbons on the other phenyl rings (Ph1 and Ph2) part of the dPPO arms in PPT (see Table 1) confirming the  $\pi$ -conjugation breaking between the core and the peripheral dPPO moieties already predicted for PPT in a previous work.<sup>15</sup>

However, when comparing the same relative  $C_i$  positions of Ph1 and Ph2 in TPPO and in PPT, it is worth noting that in PPT they have a lower calculated C 1s IP indicating a stabilization when they are part of the more complex PPT molecule.

**O 1s Core Level.** O 1s core level spectra of TPPO and PPT, taken at a photon energy of 628 eV are shown in Figure 4a,b



**Figure 4.** O 1s XP spectra of TPPO (a), PPT (b), and their comparison in the upper panel. Experimental data (dots) are shown together with the results of the total fitting procedure (black line). The vertical bars are the theoretical IPs shifted by  $-0.12$  and  $-0.16$  eV for TPPO and PPT, respectively.

respectively, whereas the comparison of their experimental lineshapes is reported in the upper panel. The spectra show for both molecules a quite broad peak centered at 535.53 eV for TPPO and at 535.44 eV for PPT. We observe a broadening of the O 1s signal for PPT (fwhm = 0.820 eV) with respect to TPPO (fwhm = 0.741 eV) as well as a very small shift toward lower BEs (of about 90 meV), analogous to the stabilization of the C 1s XPS lines when passing from the building blocks to the more complex PPT molecule.

The experimental O 1s XP spectrum of TPPO is well reproduced by the fit considering the calculated IP just shifted  $-0.12$  eV. Table 2 reports all the experimental, fitted, and theoretical data for the O 1s XPS results of TPPO and PPT. The calculations split the O 1s signal of PPT into two very close components (the bars in Figure 4b): one line (O2) with the same IP value found in TPPO and the other line (O1),

lower in energy by 80 meV. This well describes the experimental broadening of the PPT O 1s XPS line with respect to the TPPO results. In Figure 4a,b, the experimental versus the total fitted lineshapes are shown together with the calculated IPs reported as bars representing the theoretical O 1s IP of TPPO and the IPs of the O1 and O2 components of PPT (see the molecular inset in the spectrum), respectively.

The theoretical O 1s split reflects the slightly different dihedral  $O=P-C-C$  angle of the two dPPO arms of PPT (see SI Table S1b) and the consequent nonequivalence of the O1 and O2 atoms. Indeed, also previous investigations have found that TPPO and PPT are not planar and characterized by different torsional  $O=P-C-C$  angles generated by the rotation of the phenyls around the three  $P-C$  bonds.<sup>42,30</sup> This has been verified in the optimized structures for PPT and TPPO presented here in Figure 2a,b and confirmed to be the reason for the splitting into O1 and O2 components in our DFT predicted O 1s spectrum for PPT.

## EXPERIMENTAL VERSUS THEORETICAL NEXAFS

**C 1s NEXAFS.** In the following discussion, the molecular orbitals (MOs) involved in the most intense transitions of TPPO and PPT are considered and their composition analyzed in terms of Mulliken population of atomic orbitals (AOs) that are centered on each specific atom. Figure 5 shows the comparison between the experimental and theoretical C 1s NEXAFS spectra of TPPO and PPT. Also, the results for DBT from our previous study<sup>16</sup> have been included for illustrating the trend of the spectral features in going from the DBT and TPPO building blocks to the PPT system.

As seen in Figure 5, the calculated spectra agree quite well with the experiments being all the main experimental features correctly predicted by the theory, especially in the low photon energy range of the spectra (bands A and B). The main peak assignments for TPPO, PPT, and DBT systems are reported in Table 3. More detailed assignments of all nonequivalent  $C_i$  spectral features can be found in the SI Figures S3 and S4 and corresponding Tables S4 and S5, while for the complete DBT C 1s NEXAFS bands assignments we refer to our previously published study.<sup>16</sup>

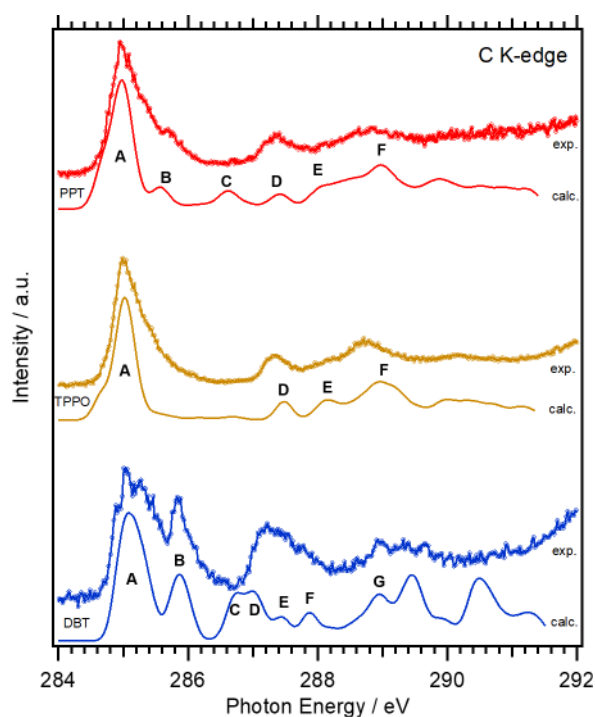
**TPPO C 1s NEXAFS.** The experimental C 1s NEXAFS spectrum of TPPO is characterized by a sharp and intense resonance (Figure 5, peak A) at 284.96 eV similarly to PPT. Going on the higher photon energy side, the TPPO spectrum does not show any band corresponding to features B and C visible in both PPT and DBT spectra, being typical for thiophene-containing molecules. At higher photon energies, bands D, E, and F are at 287.30, 288.24, and 288.71 eV, respectively, centered in the TPPO spectrum. The calculated spectrum correctly reproduces the experimental features, in particular their energy separations as well as the intensity distribution. For a detailed theoretical assignment of the

**Table 2.** O 1s Experimental (BEs) and Theoretical (IPs) Results of PPT and TPPO<sup>a</sup>

TPPO				PPT			
O 1s peak	IP <sup>(ΔSCF)</sup> <sup>b</sup>	BE	fwhm	O 1s Peaks	IP <sup>(ΔSCF)</sup> <sup>c</sup>	BE	fwhm
O	535.53	535.53	0.741	O1	535.41	535.44	0.820
				O2	535.49		

<sup>a</sup>The BEs have been obtained by fitting the experimental O 1s XP spectra. All values are expressed in eV. <sup>b</sup>Calculated IPs shifted by  $-0.12$  eV (TPPO). <sup>c</sup>Calculated IPs shifted by  $-0.16$  eV (PPT) in order to match the experimental peaks.





**Figure 5.** Comparison of the experimental (exp.) and calculated (calc.) C 1s NEXAFS spectra of PPT, TPPO, and DBT molecules.<sup>16</sup> The main calculated features are indicated with capital letters. Each calculated total spectrum has been obtained from the convolution of the partial  $C_i$  contributions (shown in the SI) with Gaussians with  $\text{fwhm} = 0.30$  eV for TPPO and DBT and  $\text{fwhm} = 0.40$  eV for PPT. The calculated spectra have been shifted by  $-0.58$  eV (PPT),  $-0.63$  eV (TPPO), and  $-0.40$  eV (DBT)<sup>16</sup> in order to match the first experimental peak.

experimental features of the C 1s NEXAFS spectrum of TPPO, we refer to Figure S3 and Table S4 of the SI.

The inspection of the site-resolved partial spectra reported in Figure S3 reveals that all the features of TPPO spectrum are contributed by all  $C_i$  atoms of the phenyl rings in very similar extent in terms of both transition energies and oscillator strengths. In particular, the calculated peak A derives from the C 1s  $\rightarrow$  LUMO transitions from all the  $C_i$  sites. As it is shown in the sketch of Figure S3, the ground state GS-LUMO is a  $\pi^*$  orbital delocalized on the entire molecule highlighting the aromatic character of TPPO. However, in the presence of a core-hole at the  $C_3$  site, a drastic charge density redistribution from Ph2 and Ph3 rings toward the ring (Ph1) containing the core-hole site is observed. The lower energy shoulder of peak A is attributed to the LUMO transition from the  $C_{\text{ipso}}$  sites and occurs at about 300 meV lower than the other C 1s  $\rightarrow$  LUMO

transitions. This shift reflects the XPS IP chemical shift observed for the  $C_{\text{ipso}}$  sites in the case of TPPO.

The two broad spectral features (peaks D and E in Figure 5) in the 287.30–288.70 eV energy range, both have a mixed diffuse Rydberg and valence  $\sigma^*(\text{C-H})$  character. Transitions from all the  $C_i$  sites (Figure S3) of each phenyl ring contribute to these bands, except from  $C_{\text{ipso}}$  and  $C_{\text{ortho}}$  for which negligible intensity is calculated; a very similar trend among the three phenyl rings is present. In going toward the ionization thresholds, a large feature is present (labeled F and centered experimentally at 288.71 eV) which arises from a manifold of transitions from all the  $C_i$  sites toward virtual MOs of mixed valence  $\pi^*(\text{C=C})$ /Rydberg nature.

**PPT C 1s NEXAFS.** The experimental C 1s NEXAFS spectrum of PPT in Figure 5 is also dominated by an intense low energy peak (A) at 284.95 eV with a higher energy feature (B) around 285.69 eV. This main feature is followed by less intense and broader structures (peaks C to F) falling in the energy region between 286.67 and 288.64 eV, that is, below the ionization threshold. A good agreement between the experimental and the theoretical NEXAFS spectra is also found for the PPT molecule, and a detailed assignment of the spectral features is reported in Table S5 and in Figure S4 of the SI which highlight the contributions of the three phenyl rings to the calculated NEXAFS spectrum.

The most striking observation from the theoretical analysis (reported in SI) is that the contributions of the Ph3 phenyl ring to the total spectrum differ significantly from the contributions of the other rings being that the Ph3 ring is part of the DBT core of PPT. The contributions of Ph1 and Ph2 rings instead appear similar resembling what we already found for the XPS spectrum. Table S5 also helps to assign peak A to the LUMO transitions from all the  $C_i$  sites except  $C_{\text{para}}$  (bonded to the S atom) of Ph3 which contributes to feature B of the spectrum. The shift to higher energy of the Ph3  $C_{\text{para}}$  excitation is in agreement with the energy position of the relative calculated IP of the Ph3 ring. In addition, we observe that the LUMO excitation energies of the other  $C_i$  sites of Ph3 are slightly lower than those of Ph1 and Ph2 still in substantial agreement with the trend observed for the IPs of the ortho/ortho' and meta/meta' positions of Ph3. Furthermore, the GS-LUMO orbital, shown in the right side panel of Figure S4, has a  $\pi^*(\text{C=C})$  character of the phenyl rings and is delocalized on the entire DBT molecular moiety. Moreover, as already found for TPPO, for PPT when creating a core-hole at sites belonging to Ph1 and Ph2 (i.e., at  $C_{29}$  and  $C_{30}$ , respectively), the charge density is also redistributed and localized at the phenyl ring to which the core-hole site belongs. On the contrary, when we create a core-hole on the  $C_i$  site of Ph3 ring (i.e., on  $C_{32}$  and  $C_{36}$ ), the charge density does not shift

**Table 3.** Assignments of the Main Calculated Features of the C 1s Spectra of TPPO, PPT, and DBT Molecular Systems<sup>a</sup>

peaks	TPPO		PPT		DBT	
	$E$ (exp.)	assignment	$E$ (exp.)	assignment	$E$ (exp.)	assignment
A	284.96	LUMO: $1\pi^*(\text{C=C})$	284.95	LUMO: $1\pi^*(\text{C=C})$	284.95–285.57	LUMO: $1\pi^*(\text{C=C})$
B			285.69	$1\pi^*(\text{C=C})$	285.89	$1\pi^*(\text{C=C})$
C			286.67	$\sigma^*(\text{C-S})$ $\sigma^*(\text{C-H})$ Rydberg mixed	287.15	$\sigma^*(\text{C-S})$
D	287.30	$\sigma^*(\text{C-H})$ –Rydberg mixed	287.34			$2\pi^*$
E	288.24		288.22		287.49	$\pi^*(\text{C=C})$ –Rydberg mixed
F	288.71	$\pi^*(\text{C=C})$ Rydberg mixed	288.64	$\pi^*(\text{C=C})$ –Rydberg mixed Rydberg	288.07	Rydberg

<sup>a</sup>All values are in eV.

toward Ph1 and Ph2 but it still remains confined inside the Ph3 ring. A difference about the charge density distribution can be discussed as follows: in the TP-LUMO with a core-hole on C<sub>36</sub> site, a higher conjugation across the whole DBT core system is found, while a lower conjugation extension over the DBT core system is observed creating a core-hole on C<sub>32</sub> site (i.e., TP-LUMO with core-hole on C<sub>32</sub> site).

Peak C is still associated with several Ph3 transitions, the most intense of which involves a  $\sigma^*(\text{C-S})$  orbital, while peak D is derived from the C<sub>i</sub> sites transitions of Ph1 and Ph2 (see Table S5). The two following structures (labeled E and F) are characterized by a superposition of core excitations to virtual MOs of mixed valence and Rydberg character starting from the C atoms of all the three phenyl rings.

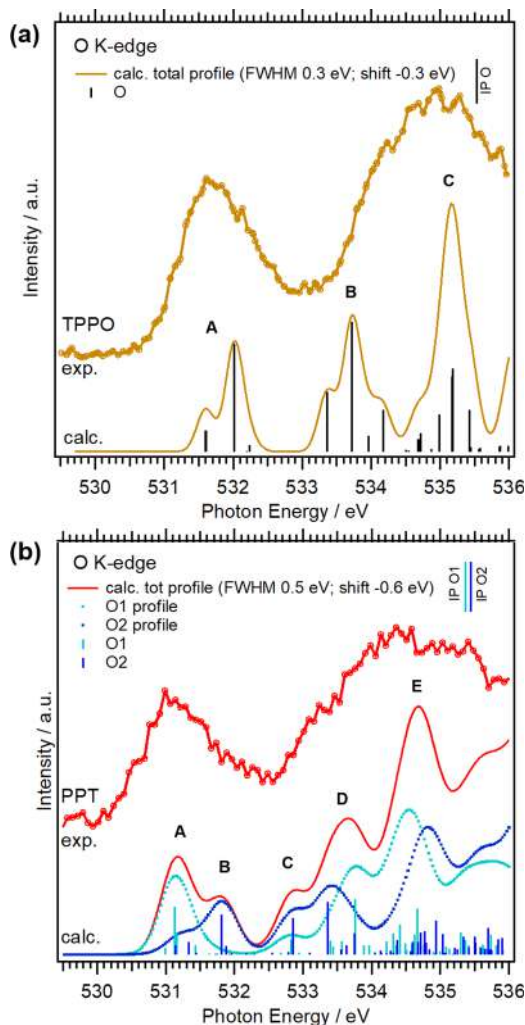
**O 1s NEXAFS.** O 1s NEXAFS spectra of the TPPO and PPT molecules are shown in Figure 6a,b, respectively, together with theoretical data and band assignments. The theoretical

results, in terms of O 1s core excitation energies and oscillator strengths, are reported in Table 4 (TPPO molecule) and in

**Table 4. Peak Assignments for the O 1s NEXAFS Spectrum of TPPO**

peak	$E$ calc. shifted <sup>a</sup> (eV)	$f \times 10^{2b}$	assignments	$E$ exp. (eV)
A	531.60	0.071	$\pi^*(\text{C}=\text{C})$	
	532.02	0.36	$\pi^*(\text{C}=\text{C}) + \text{P } 3p$	
B	533.36	0.20	Rydberg	533.20
	533.72	0.44	Rydberg/ $\pi^*(\text{C}=\text{C})$	
	534.18	0.14	Rydberg/ $\sigma^*(\text{P-O})$	
C	534.99	0.12	Rydberg	535.00
	534.99	0.13		
	535.17	0.18		
	535.18	0.25	mixed valence [ $\pi^*(\text{C}=\text{C}) +$	
	535.18	0.28	$\pi^*(\text{P-O})$ ]/Rydberg	

<sup>a</sup>Calculated excitation energies shifted by  $-0.3$  eV to match the first experimental peak. <sup>b</sup>Only transitions with  $f \times 10^2 \geq 0.07$  are reported.



**Figure 6.** Comparison between the O 1s NEXAFS experimental (exp.) and calculated (calc.) spectra of (a) TPPO and (b) PPT. The partial O<sub>i</sub> contributions (O1 in light blue and O2 dark blue) are also shown in the PPT spectrum (bottom of panel b). The  $\Delta\text{KS}$  O 1s IPs are reported as vertical bars (top right side of the panels). The theoretical profiles have been convoluted with Gaussians of fwhm = 0.3 eV (TPPO) and fwhm = 0.5 eV (PPT). The calculated peaks have been shifted by  $-0.3$  eV (TPPO) and  $-0.6$  eV (PPT) in order to match the first experimental peak.

Table 5 (PPT molecule). The experimental O 1s NEXAFS spectra of the TPPO and PPT are very similar, with two broad features around 531.20 and 534.70 eV, respectively.

**TPPO O 1s NEXAFS.** The TPPO O 1s calculated NEXAFS spectrum is characterized by two broad features: a lower energy band (A) located around 531.60 eV and well separated from a second series of intense features (band B) in the range 533.40–534.20 eV.

In Table 4 we see that the oscillator strength for transitions at the oxygen K threshold are very weak, particularly in the lower energy range of the spectrum. This underlines that the atomic contribution of O 2p to virtual valence levels is very small. This is in accordance with the strong polarization of the P—O bond and the large involvement of O 2p atomic components in the occupied MOs with a consequent very small contribution to the virtual valence MOs. In particular, the main contribution to peak A is from the transition at 532.02 eV toward a virtual MO with main  $\pi^*(\text{C}=\text{C})$  character and only minimal participation of O 2p components. Transition to LUMO (calculated at 531.60 eV) has a negligible theoretical intensity that is equal to one-fifth of the previous one and is presumably hidden in the main structure mainly due to the low signal-to-noise ratio of the measurement, due precisely to the small cross section of the transition under consideration.

Final states of diffuse nature characterize both B and C features, where also  $\pi^*(\text{C}=\text{C})$  and  $\sigma^*(\text{P-O})$  components are discernible. At higher energy (calculated at 535.18 eV), near the O 1s ionization threshold, two transitions appear toward MOs of mixed valence/Rydberg character with a valence contribution also from the  $\pi^*(\text{P-O})$  component. The theoretical results also indicate that the O 1s  $\rightarrow \sigma^*(\text{P-O})$  transitions are lower in energy than the O 1s  $\rightarrow \pi^*(\text{P-O})$  transitions (see Table 4) in agreement with the stabilization of the  $\pi$ -symmetry interaction in the P—O bond with respect to the  $\sigma$  one in the occupied MO of TPPO. This will be discussed in more detail when describing the P—O bond.

In general, for the O 1s NEXAFS of the TPPO molecule the agreement between experiment and theory is less satisfactory than that found for the C 1s results, in particular the energy separation between the A and B structures is probably underestimated in the calculations, although the low resolution

Table 5. Peak Assignments for the O 1s NEXAFS Spectrum of PPT

peak	site	$E$ calc. shifted <sup>a</sup> (eV)	$f \times 10^{2b}$	assignments	$E$ exp. (eV)
A	O <sub>1</sub>	530.99	$4.33 \times 10^{-2}$	$\pi^*(C=C)+P$ 3p-LUMO	531.20
	O <sub>1</sub>	531.13	0.32	$\pi^*(C=C)+(P,S)$ 3p	
	O <sub>2</sub>	531.14	$6.08 \times 10^{-2}$	$\pi^*(C=C)+P$ 3p-LUMO	
	O <sub>1</sub>	531.16	0.14	$\pi^*(C=C)+P$ 3p	
B	O <sub>2</sub>	531.81	0.27	$\pi^*(C=C)+P$ 3p	531.94
C	O <sub>1</sub>	532.82	0.10	$\pi^*(C=C)+(P,S)$ 3p	532.92
	O <sub>2</sub>	532.86	0.24	$\pi^*(C=C)+P$ 3p	
D	O <sub>2</sub>	533.36	0.35	$\pi^*(C=C)+P$ 3p	533.67
	O <sub>1</sub>	533.39	0.10	Rydberg	
	O <sub>2</sub>	533.75	0.14	Rydberg/ $\sigma^*(P=O)$	
	O <sub>1</sub>	533.76	0.37		
E	O <sub>1</sub>	534.31	0.14	Rydberg	534.70
	O <sub>2</sub>	534.37	0.10	Rydberg	
	O <sub>1</sub>	534.41	0.20	Rydberg/ $\pi^*(P=O)$	
	O <sub>1</sub>	534.62	0.13	Rydberg	
	O <sub>1</sub>	534.66	0.30	mixed valence [ $\pi^*(C=C) + \pi^*(P=O)$ ]/Rydberg	
	O <sub>2</sub>	534.72	0.15	Rydberg	
	O <sub>2</sub>	534.77	0.15	Rydberg	
	O <sub>2</sub>	534.93	0.22	mixed valence [ $\pi^*(C=C) + \pi^*(P=O)$ ]/Rydberg	

<sup>a</sup>Calculated excitation energies shifted by  $-0.6$  eV to match the first experimental peak. <sup>b</sup>Only transitions with  $f \times 10^2 \geq 0.07$  are reported; the lower intensity O1 and O2 to LUMO transitions are also included.

and the poor signal-to-noise ratio of the experiment prevents a more detailed comparison of the results.

**PPT O 1s NEXAFS.** In the case of PPT, the theoretical results reported in Figure 6b allow us to distinguish the contributions to the O 1s NEXAFS spectrum of the two nonequivalent O (O1 and O2) atoms. The oscillator strengths of the O 1s transitions are still very low confirming, also in the PPT case, the scarce O 2p participation to the virtual MOs of these molecules. This is particularly evident in the lower energy region of the PPT spectrum (bands A and B) where the transitions toward MOs with predominant  $\pi^*(C=C)$  character are present at experimental energy of 531.20 and 531.94 eV, respectively. The splitting of the lower energy band into two calculated components reflects the chemical shift proposed for O1 and O2 atoms when discussing the XPS results.

For the PPT molecule, the theory predicts that band A is mainly due to the O1 transitions to MOs of predominant  $\pi^*(C=C)$  character, while band B mainly arises from the corresponding O2 transitions, although transitions from O2 also contribute to band A. The different intensity distributions among the lines of A and B bands reflect the different geometrical environment of the two O atoms, also confirmed by the calculated geometrical parameters (see Table S1b).

Transitions toward MOs with predominant  $\pi^*(C=C)$  components also give rise in the PPT spectrum to feature C (exp. at 532.92 eV), while peak D (exp. at 533.67 eV) derives from a manifold of transitions of different nature. Among these, also transitions to MOs with valence  $\sigma^*(P=O)$  component are present (calc. at 533.75 and 533.76 eV), in an energy region at about 2.8 eV from the LUMO transition, therefore in line with what also found for the O K-edge NEXAFS spectrum of TPPO.

Also the higher energy E structure for PPT (exp. at 534.70 eV) derives its intensity from a manifold of transitions where the final MOs have either mixed valence/Rydberg or pure Rydberg character. The valence character of the most intense transitions involves both  $\pi^*(C=C)$  and  $\pi^*(P=O)$  compo-

nents, in line with the spectral trend found for band C of TPPO molecule falling in a similar energy range (around 535 eV).

The description of the P—O bond in phosphine oxides has raised great interest for many years. It has been extensively reviewed by Gilheany<sup>43</sup> who observed that both experiment and *ab initio* calculations generally show that the P—O bond is strong, polar, and as short as conventional P—O double bonds. The strong polarization of the P—O bond is also supported by our Bader's population analysis. Indeed, as reported in Table S3 of the SI file, the P atom carries a partial positive charge and the electronegative O atom carries a partial negative one, causing a depletion of electron density away from P which is true both for TPPO and PPT molecular systems. Moreover, the O 1s NEXAFS spectra calculations have provided very low intensities for the lower-energy transitions, pointing toward a small participation of O 2p AOs to virtual valence MOs and consequently to a greater extent into the occupied MOs.

Even though it is well-known that the P—O bond is multiple and highly polar, there is still strong disagreement about the exact electron distribution on it.

In this respect, there are three viewpoints: the first one, which is the more traditional, considers the formation of the P—O bond as one  $\sigma$  from P to O, through participation of the lone pair on P, and two  $\pi$ -back-bonds from O lone-pairs to P acceptor orbitals of  $e$  type and  $\sigma^*$  symmetry (negative hyperconjugation)<sup>44</sup> giving to the P—O bond a partial triple-bond character.<sup>45</sup> The mechanism of negative hyperconjugation is also suggested by our O 1s NEXAFS calculations which show that transitions toward  $\sigma^*(P=O)$  states occur at lower energies with respect to those toward  $\pi^*(P=O)$  states.

An alternative possibility takes into account one  $\sigma$ -bond and three  $\pi$ -back-bonds from the O lone pairs. Finally, the last viewpoint agrees with three bent multiple bonds ( $\Omega$ -bonds) strongly polarized toward the O atom, so that the P—O bond is a formal triple bond with the three curved regions of electron density arranged between P and O in a symmetrical fashion at

120° to each other in Newmann projection along the P—O coordinate. The  $\Omega$ -bond description is also provided by the Boys localization procedure for phosphine oxides in SCF-HF/LCAO-MO calculations.<sup>46–48</sup> This last view is however not supported by our NEXAFS results.

## CONCLUSIONS

The main goal of the present study is the evaluation of the impact of the single building blocks (DBT and TPPO) on the electronic structure of the whole PPT ambipolar system through the C 1s and O 1s core levels analysis by XPS and NEXAFS synchrotron-based spectroscopies.

The C 1s XPS trends of C<sub>i</sub> sites in Ph1 and Ph2 rings are quite similar for TPPO and PPT molecules. This means that the Ph1 and Ph2 rings, representing the dPPO arms in PPT, conserve the typical TPPO electronic characteristics also when part of the PPT molecule and the presence of the DBT moiety does not influence their IPs significantly.

The theoretical IPs of the C<sub>i</sub> of the Ph3 ring, part of the PPT core, are instead lower than those of the corresponding carbons on the other phenyl rings (Ph1 and Ph2). This theoretical result could be considered as a proof of a higher electronic stabilization, ascribed to an increased available charge within the conjugated system of the central DBT moiety to which the Ph3 belongs. This also confirms the conjugation breaking in correspondence to the P—C<sub>32</sub> bond, being that the Ph3 ring is an integral part of the DBT core in PPT. Furthermore, as a consequence of different geometric surroundings (see torsional angles reported in detail in Table S1 of SI), the two oxygen atoms of PPT are predicted to be inequivalent (O1 and O2). This is also confirmed by the good agreement between the experimental and theoretical O 1s XP and O K-edge NEXAFS spectra obtained by using two theoretical profiles for O1 and O2, respectively.

The main achievement of our NEXAFS O 1s calculations provides, both for TPPO and PPT molecules, that the O 2p atomic contribution to the valence virtual levels is very small. This is in accordance with the strong polarization of the P—O bond and the large involvement of O 2p atomic components in the occupied MOs with a consequent very small contribution to the virtual valence MOs. In this respect, an investigation at the NEXAFS P K-edge would be also useful, thus providing specific information on the weight of the P 2p AOs contributions. Moreover, from the calculations related to the NEXAFS O 1s data analysis, it has been highlighted that the O 1s  $\rightarrow \sigma^*(\text{P—O})$  transitions are lower in energy than the O 1s  $\rightarrow \pi^*(\text{P—O})$  transitions in agreement with the stabilization of the  $\pi$ -symmetry interaction in the P—O bond with respect to the  $\sigma$ -one in the occupied MOs of TPPO.

In conclusion, the results presented in this study show that the conjugation breaking acting across the P—C<sub>32</sub> bond connecting the DBT core and the dPPO arms (already predicted in previous works)<sup>15</sup> is one of the peculiar characteristics of the PPT molecular system. It is in fact known<sup>15</sup> that the phosphorus center in PPT behaves as a  $\pi$ -conjugation blocker among the central DBT electron-donor moiety and the peripheral dPPO electron-acceptor counterparts, making the PPT molecular system a useful ambipolar material. The conjugation breaking guarantees that the central DBT core of PPT maintains its triplet energy, which is important in OLEDs applications. This allows one to preserve the photophysical properties of DBT and to exploit them after

incorporating this volatile moiety into a vacuum sublimable material such as PPT.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcc.0c01764>.

Geometry optimizations results and the complete assignments of the intensities contribution to the C K-edge NEXAFS spectra for TPPO and PPT molecules (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

**Ambra Guarnaccio** — ISM-CNR, Istituto di Struttura della Materia, Tito Scalo (Pz) and Trieste, Italy; [orcid.org/0000-0002-7927-5845](https://orcid.org/0000-0002-7927-5845); Email: [ambra.guarnaccio@ism.cnr.it](mailto:ambra.guarnaccio@ism.cnr.it)

**Giovanna Fronzoni** — Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy; Email: [fronzoni@units.it](mailto:fronzoni@units.it)

### Authors

**Teng Zhang** — School of Information and Electronics, MIT Key Laboratory for Low-Dimensional Quantum Structure and Devices, Beijing Institute of Technology (BIT), Beijing 100081, China; Department of Physics and Astronomy, Uppsala University, SE-751 20 Uppsala, Sweden; [orcid.org/0000-0001-8739-7773](https://orcid.org/0000-0001-8739-7773)

**Cesare Grazioli** — IOM-CNR, Laboratorio TASC, Sincrotrone Trieste, I-34149 Trieste, Basovizza, Italy

**Fredrik O. L. Johansson** — Department of Physics and Astronomy, Uppsala University, SE-751 20 Uppsala, Sweden; [orcid.org/0000-0002-6471-1093](https://orcid.org/0000-0002-6471-1093)

**Marcello Coreno** — ISM-CNR, Istituto di Struttura della Materia, Tito Scalo (Pz) and Trieste, Italy

**Monica de Simone** — IOM-CNR, Laboratorio TASC, Sincrotrone Trieste, I-34149 Trieste, Basovizza, Italy

**Daniele Toffoli** — Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy

**Elisa Bernes** — Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy

**Carla Puglia** — Department of Physics and Astronomy, Uppsala University, SE-751 20 Uppsala, Sweden; [orcid.org/0000-0001-6840-1570](https://orcid.org/0000-0001-6840-1570)

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpcc.0c01764>

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors acknowledge the EU CERIC—ERIC Consortium for the access to experimental facilities and financial support. The authors thank the Carl Trygger Foundation for financial support and for making available the VG-Scienta SES-200 photoelectron analyzer at the Gas Phase beamline, Elettra, Italy. T.Z. is grateful for the financial support from the NSFC (Grants 61901038 and 61971035) and the support from the Beijing Institute of Technology Research Fund Program for Young Scholars. F.J. acknowledges the support from the Swedish Research Council (Grant 2014-6463) and Marie Skłodowska Curie Actions (Cofund, Project INCA 600398).

Computational research has been supported by Finanziamento per ricerca di ateneo, FRA 2015 and FRA 2016 of the Università degli Studi di Trieste.

## REFERENCES

- (1) Roncali, J. Synthetic principles for bandgap control in linear  $\pi$ -conjugated systems. *Chem. Rev.* **1997**, *97*, 173–205.
- (2) Tsai, Y. S.; Hong, L. A.; Juang, F. S.; Chen, C. Y. Blue and white phosphorescent organic light emitting diode performance improvement by confining electrons and holes inside double emitting layers. *J. Lumin.* **2014**, *153*, 312–316.
- (3) Chapran, M.; Angioni, E.; Findlay, N. J.; Breig, B.; Cherpak, V.; Stakhira, P.; Tuttle, T.; Volyniuk, D.; Grazulevicius, J. V.; Nastishin, Y. A.; et al. An ambipolar BODIPY derivative for a white exciplex OLED and cholesteric liquid crystal laser toward multifunctional devices. *ACS Appl. Mater. Interfaces* **2017**, *9*, 4750–4757.
- (4) Nishide, J. I.; Nakanotani, H.; Hiraga, Y.; Adachi, C. High-efficiency white organic light-emitting diodes using thermally activated delayed fluorescence. *Appl. Phys. Lett.* **2014**, *104*, 233304–5.
- (5) Sun, J. W.; Kim, K. H.; Moon, C. K.; Lee, J. H.; Kim, J. J. Highly efficient sky-blue fluorescent organic light emitting diode based on mixed cohost system for thermally activated delayed fluorescence emitter (2CzPN). *ACS Appl. Mater. Interfaces* **2016**, *8*, 9806–9810.
- (6) Goushi, K.; Yoshida, K.; Sato, K.; Adachi, C. Organic light-emitting diodes employing efficient reverse intersystem crossing for triplet-to-singlet state conversion. *Nat. Photonics* **2012**, *6*, 253–258.
- (7) Goushi, K.; Adachi, C. Efficient organic light-emitting diodes through up-conversion from triplet to singlet excited states of exciplexes. *Appl. Phys. Lett.* **2012**, *101*, 023306–4.
- (8) Su, S. J.; Sasabe, H.; Takeda, T.; Kido, J. Pyridine-Containing Bipolar Host Materials for Highly Efficient Blue Phosphorescent OLEDs. *Chem. Mater.* **2008**, *20*, 1691–1693.
- (9) Hsu, F. M.; Chien, C. H.; Shih, P. I.; Shu, C. F. Phosphine-oxide-containing bipolar host material for blue electrophosphorescent devices. *Chem. Mater.* **2009**, *21*, 1017–1022.
- (10) Sapochak, L. S.; Padmaperuma, A. B.; Vecchi, P. A.; Cai, X.; Burrows, P. E. Designing organic phosphine oxide host materials using heteroarmatic building blocks: inductive effects on electroluminescence. *Proc. SPIE* **2007**, *6655*, 665506.
- (11) Burrows, P. E.; Padmaperuma, A.; Sapochak, L. S.; Djurovich, P.; Thompson, M. E. Ultraviolet electroluminescence and blue-green phosphorescence using an organic diphosphine oxide charge transporting layer. *Appl. Phys. Lett.* **2006**, *88*, 183503–3.
- (12) Padmaperuma, A. B.; Sapochak, L. S.; Burrows, P. E. New charge transporting host material for short wavelength organic electrophosphorescence: 2,7-Bis(diphenylphosphine oxide)-9,9-dimethylfluorene. *Chem. Mater.* **2006**, *18*, 2389–2396.
- (13) Cai, X.; Padmaperuma, A. B.; Sapochak, L. S.; Vecchi, P. A.; Burrows, P. E. Electron and hole transport in a wide bandgap organic phosphine oxide for blue electrophosphorescence. *Appl. Phys. Lett.* **2008**, *92*, 083308–3.
- (14) Polikarpov, E.; Swensen, J. S.; Chopra, N.; So, F.; Padmaperuma, A. B. An ambipolar phosphine oxide-based host for high power efficiency blue phosphorescent organic light emitting devices. *Appl. Phys. Lett.* **2009**, *94*, 223304–3.
- (15) Kim, D.; Salman, S.; Coropceanu, V.; Salomon, E.; Padmaperuma, A. B.; Sapochak, L. S.; Kahn, A.; Bredas, J. L. Phosphine oxide derivatives as hosts for blue phosphors: a joint theoretical and experimental study of their electronic structure. *Chem. Mater.* **2010**, *22*, 247–254.
- (16) Toffoli, D.; Guarnaccio, A.; Grazioli, C.; Zhang, T.; Johansson, F.; de Simone, M.; Coreno, M.; Santagata, A.; D'Auria, M.; Puglia, C.; et al. Electronic structure characterization of a thiophene benzoannulated series of common building blocks for donor and acceptor compounds studied by gas phase photoelectron and photoabsorption synchrotron spectroscopies. *J. Phys. Chem. A* **2018**, *122*, 8745–8761.
- (17) Grazioli, C.; Baseggio, O.; Stener, M.; Fronzoni, G.; de Simone, M.; Coreno, M.; Guarnaccio, A.; Santagata, A.; D'Auria, M. Study of the electronic structure of short chain oligothiophenes. *J. Chem. Phys.* **2017**, *146*, 054303–9.
- (18) Baseggio, O.; Toffoli, D.; Stener, M.; Fronzoni, G.; de Simone, M.; Grazioli, C.; Coreno, M.; Guarnaccio, A.; Santagata, A.; D'Auria, M. S<sub>2p</sub> core level spectroscopy of short chain oligothiophenes. *J. Chem. Phys.* **2017**, *147*, 244301–13.
- (19) Blyth, R.; Delaunay, R.; Zitnik, M.; Krempasky, J.; Krempaska, R.; Slezak, J.; Prince, K. C.; Richter, R.; Vondracek, M.; Camilloni, R.; et al. The high resolution Gas Phase Photoemission beamline, Elettra. *J. Electron Spectrosc. Relat. Phenom.* **1999**, *101*, 959–964.
- (20) Mårtensson, N.; Baltzer, P.; Brühwiler, P. A.; Forsell, J. O.; Nilsson, A.; Stenborg, A.; Wannberg, B. A very high resolution electron spectrometer. *J. Electron Spectrosc. Relat. Phenom.* **1994**, *70*, 117–128.
- (21) Myrseth, V.; Bozek, J.; Kukk, E.; Sæthre, L.; Thomas, T. Adiabatic and vertical carbon 1s ionization energies in representative small molecules. *J. Electron Spectrosc. Relat. Phenom.* **2002**, *122*, 57–63.
- (22) Johansson, G.; Hedman, J.; Berndtsson, A.; Klasson, M.; Nilsson, R. Calibration of electron spectra. *J. Electron Spectrosc. Relat. Phenom.* **1973**, *2*, 295–317.
- (23) Schmid, M.; Steinrück, H.; Gottfried, J. M. A new asymmetric Pseudo-Voigt function for more efficient fitting of XPS lines. *Surf. Interface Anal.* **2014**, *46*, 505–511.
- (24) Tronc, M.; King, G. C.; Read, F. H. Nitrogen K-shell excitation in N<sub>2</sub>, NO and N<sub>2</sub>O by high-resolution electron energy-loss spectroscopy. *J. Phys. B: At. Mol. Phys.* **1980**, *13*, 999–1008.
- (25) 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.
- (26) Parr, R. G.; Yang, W. *Density Functional Theory of atoms and molecules*; Oxford University Press: New York, 1989.
- (27) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- (28) Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. Towards an order-N DFT method. *Theor. Chem. Acc.* **1998**, *99*, 391–403.
- (29) 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.
- (30) Bandoli, G.; Bortolozzo, G.; Clemente, D. A.; Croatto, U.; Panattoni, C. Crystal and molecular structure of triphenylphosphine oxide. *J. Chem. Soc. A* **1970**, 2778–2780.
- (31) Al-Farhan, K. A. Crystal structure of triphenylphosphine oxide. *J. Crystallogr. Spectrosc. Res.* **1992**, *22*, 687–689.
- (32) Diop, C. A. K.; Touré, A.; Diop, L.; Welter, R. 2-Benzoylbenzoato- $\kappa$ O)triphenyl(triphenylphosphineoxide- $\kappa$ O)tin(IV). *Acta Crystallogr., Sect. E: Struct. Rep. Online* **2006**, *E62*, m3338–m3340.
- (33) De Silva, N.; Zahariev, F.; Hay, B. P.; Gordon, M. S.; Windus, T. L. Conformations of organophosphine oxides. *J. Phys. Chem. A* **2015**, *119*, 8765–8773.
- (34) 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.
- (35) Rodriguez, J. I.; Köster, A. M.; Ayers, P. W.; Santos-Valle, A.; Vela, A.; Merino, G. An efficient grid-based scheme to compute QTAIM atomic properties without explicit calculation of zero-flux surfaces. *J. Comput. Chem.* **2009**, *30*, 1082–1092.
- (36) Rodriguez, J. I.; Bader, R. F. W.; Ayers, P. W.; Michel, C.; Götz, A. W.; Bo, C. A high performance grid-based algorithm for computing QTAIM properties. *Chem. Phys. Lett.* **2009**, *472*, 149–152.
- (37) Slater, J. C. Statistical exchange-correlation in the self-consistent field. *Adv. Quantum Chem.* **1972**, *6*, 1–92.
- (38) Triguero, L.; Pettersson, 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.

(39) Mijovilovich, A.; Pettersson, L. G. M.; Mangold, S.; Janousch, M.; Susini, J.; Salome, M.; de Groot, F. M. F.; Weckhuysen, B. M. The interpretation of sulfur K-Edge XANES spectra: a case study on thiophenic and aliphatic sulfur compounds. *J. Phys. Chem. A* **2009**, *113*, 2750–2756.

(40) Norman, P.; Dreuw, A. Simulating X-ray Spectroscopies and calculating core-excited states of molecules. *Chem. Rev.* **2018**, *118*, 7208–7248.

(41) Fouda, A. E. A.; Besley, N. A. Assessment of basis sets for Density Functional Theory-based calculations of core-electron spectroscopies. *Theor. Chem. Acc.* **2018**, *137*, 1–11.

(42) Al-Fahran, K. A. Crystal structure of triphenylphosphine oxide. *J. Crystallogr. Spectrosc. Res.* **1992**, *22*, 687–689.

(43) Gilheany, D. G. No d orbitals but Walsh diagrams and maybe Banana Bonds: chemical bonding in phosphines, phosphine oxides, and phosphonium ylides. *Chem. Rev.* **1994**, *94*, 1339–1374.

(44) Alabugin, I. V.; dos Passos Gomes, G.; Abdo, M. A. Hyperconjugation. *WIREs Comput. Mol. Sci.* **2019**, *9*, 1–66.

(45) Reed, A. E.; Schleyer, P. v. R. Chemical bonding in hypervalent molecules. The dominance of ionic bonding and negative hyperconjugation over d-orbital participation. *J. Am. Chem. Soc.* **1990**, *112*, 1434–1445.

(46) Schmidt, M. W.; Yabushita, S.; Gordon, M. S. Structure, bonding, and internal rotation in phosphine oxide (H<sub>3</sub>PO), hydroxyphosphine (H<sub>2</sub>POH), and hydroxyfluorophosphine (HFPOH). *J. Phys. Chem.* **1984**, *88*, 382–389.

(47) Guest, M. F.; Hillier, I. H.; Saunders, V. R. Ab initio localized orbital calculations of the bonding in phosphines and phosphine complexes. *J. Chem. Soc., Faraday Trans. 2* **1972**, *68*, 867–877.

(48) Wallmeier, H.; Kutzelnigg, W. Nature of the semipolar XO bond. Comparative ab Initio study of H<sub>3</sub>NO, H<sub>2</sub>NOH, H<sub>3</sub>PO, H<sub>2</sub>POH, H<sub>2</sub>P(O)F, H<sub>2</sub>SO, HSOH, HClO, ArO, and related molecules. *J. Am. Chem. Soc.* **1979**, *101*, 2804–2814.

**PPT Isolated Molecule and its Building Block Moieties Studied by C 1s  
and O 1s Gas Phase X-ray Photoelectron and Photoabsorption  
Spectroscopies**

*Ambra Guarnaccio<sup>a\*</sup>, Teng Zhang<sup>b,c</sup>, Cesare Grazioli<sup>d</sup>, Fredrik O. L. Johansson<sup>c</sup>,  
Marcello Coreno<sup>a</sup>, Monica de Simone<sup>d</sup>, Giovanna Fronzoni<sup>e\*</sup>, Daniele Toffoli<sup>e</sup>, Elisa  
Bernes<sup>e</sup> and Carla Puglia<sup>c</sup>*

<sup>a</sup> ISM-CNR, Istituto di Struttura della Materia - Tito Scalo (Pz) and Trieste, Italy

<sup>b</sup> School of Information and Electronics, MIIT Key Laboratory for Low-Dimensional Quantum Structure and  
Devices, Beijing Institute of Technology (BIT), Beijing 100081, China

<sup>c</sup> Department of Physics and Astronomy, Uppsala University, Box 516, SE-751 20 Uppsala, Sweden

<sup>d</sup> IOM-CNR, Laboratorio TASC, Sincrotrone Trieste, I-34149 Trieste, Basovizza, Italy

<sup>e</sup> Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy

\*Corresponding authors:

E-mail: [ambra.guarnaccio@ism.cnr.it](mailto:ambra.guarnaccio@ism.cnr.it)

E-mail: [fronzoni@units.it](mailto:fronzoni@units.it)

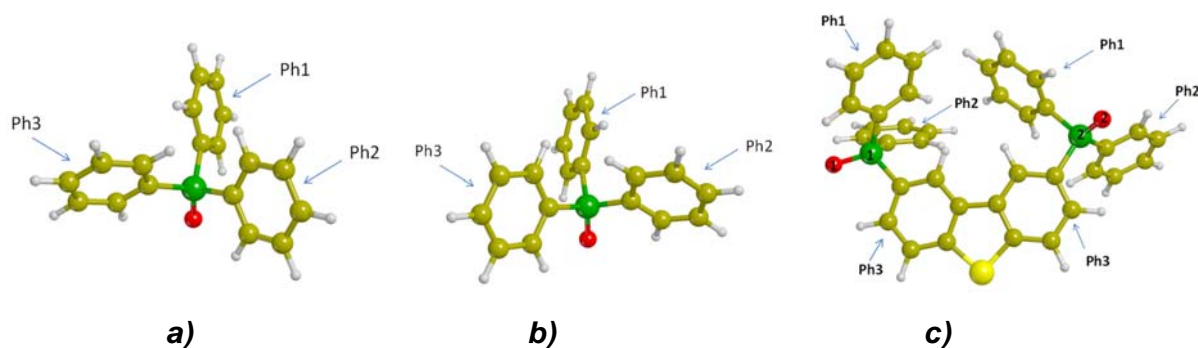
**Geometry optimization study of TPPO.....S2**

**C K-edge NEXAFS spectra and total assignments for:**

TPPO..... S7

PPT.....S9

## Geometry optimization study of TPPO



**Figure S1** - Molecular structures of the two investigated conformers of TPPO (**a**, **b**) and of PPT (**c**). The O and P atoms of PPT are also labelled.

**Table S1** - Comparison between theoretical and experimental geometrical parameters for the two investigated conformers of TPPO (**a**) and PPT (**b**). Bond lengths are expressed in Å, while angles and torsions in degrees (°).

(a)

Bond lengths	Calculated		Experimental <sup>a</sup>
	Conformer <b>a</b>	Conformer <b>b</b>	
P=O	1.491	1.491	1.479
P-C <sub>ipsoPh1</sub>	1.795	1.794	1.804
P-C <sub>ipsoPh2</sub>	1.794	1.794	1.799
P-C <sub>ipsoPh3</sub>	1.795	1.795	1.806
Angles	Calculated		Experimental <sup>a</sup>
	Conformer <b>a</b>	Conformer <b>b</b>	
O-P-C <sub>ipsoPh1</sub>	112.61	112.48	112.30
O-P-C <sub>ipsoPh2</sub>	112.55	112.58	111.90
O-P-C <sub>ipsoPh3</sub>	112.65	112.55	113.90
C <sub>ipsoPh1</sub> -P-C <sub>ipsoPh2</sub>	106.437	106.229	106.40
C <sub>ipsoPh1</sub> -P-C <sub>ipsoPh3</sub>	105.984	106.195	105.90
C <sub>ipsoPh2</sub> -P-C <sub>ipsoPh3</sub>	106.082	106.226	106.30
Torsions	Calculated		Experimental <sup>a</sup>



	Conformer <i>a</i>	Conformer <i>b</i>	
O-P-C <sub>ipsoPh1</sub> -C <sub>ortho'</sub>	-20.16	19.68	22.20
O-P-C <sub>ipsoPh2</sub> -C <sub>ortho'</sub>	-19.40	18.81	58.00
O-P-C <sub>ipsoPh3</sub> -C <sub>ortho'</sub>	159.61	-160.19	22.30

<sup>a</sup> K. A. Al-Farhan, *J. of Cryst. and Spectr. Research*, Vol. 22, 6, 1992, 687-689.

**(b)**

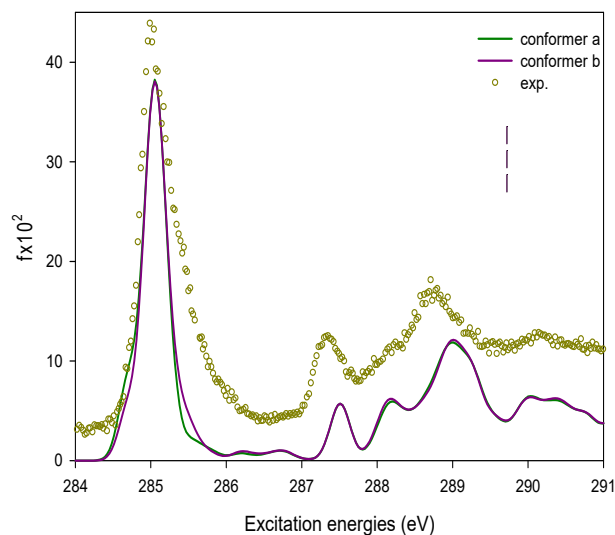
Bond lengths	Calculated	Experimental <sup>b</sup>
P1=O1	1.489	1.477
P2=O2	1.490	
P1-C <sub>ipsoPh1</sub>	1.805	1.804
P1-C <sub>ipsoPh2</sub>	1.792	1.794
P1-C <sub>ipsoPh3</sub>	1.798	1.812
P2-C <sub>ipsoPh1</sub>	1.795	1.804
P2-C <sub>ipsoPh2</sub>	1.801	1.794
P2-C <sub>ipsoPh3</sub>	1.793	1.812
Angles	Calculated	Experimental <sup>b</sup>
O1-P1-C <sub>ipsoPh1</sub>	111.8	113.4
O1-P1-C <sub>ipsoPh2</sub>	112.6	112.1
O1-P1-C <sub>ipsoPh3</sub>	117.3	111.7
O2-P2-C <sub>ipsoPh1</sub>	113.1	113.4
O2-P2-C <sub>ipsoPh2</sub>	112.9	112.1
O2-P2-C <sub>ipsoPh3</sub>	112.9	111.7
C <sub>ipsoPh1</sub> -P1-C <sub>ipsoPh2</sub>	107.6	107.3
C <sub>ipsoPh1</sub> -P1-C <sub>ipsoPh3</sub>	101.6	107.3
C <sub>ipsoPh2</sub> -P1-C <sub>ipsoPh3</sub>	105.0	104.5
C <sub>ipsoPh1</sub> -P2-C <sub>ipsoPh2</sub>	105.2	107.3
C <sub>ipsoPh1</sub> -P2-C <sub>ipsoPh3</sub>	106.5	107.3
C <sub>ipsoPh2</sub> -P2-C <sub>ipsoPh3</sub>	105.6	104.5

Torsions	Calculated
O1-P1-C <sub>ipso</sub> Ph1-C <sub>ortho</sub>	169.142
O1-P1-C <sub>ipso</sub> Ph1-C <sub>ortho'</sub>	-18.238
O1-P1-C <sub>ipso</sub> Ph2-C <sub>ortho</sub>	-19.184
O1-P1-C <sub>ipso</sub> Ph2-C <sub>ortho'</sub>	161.586
O1-P1-C <sub>ipso</sub> Ph3-C <sub>ortho</sub>	-72.852
O1-P1-C <sub>ipso</sub> Ph3-C <sub>ortho'</sub>	107.885
O2-P2-C <sub>ipso</sub> Ph1-C <sub>ortho</sub>	17.971
O2-P2-C <sub>ipso</sub> Ph1-C <sub>ortho'</sub>	-161.260
O2-P2-C <sub>ipso</sub> Ph2-C <sub>ortho</sub>	17.722
O2-P2-C <sub>ipso</sub> Ph2-C <sub>ortho'</sub>	-163.153
O2-P2-C <sub>ipso</sub> Ph3-C <sub>ortho</sub>	16.863
O2-P2-C <sub>ipso</sub> Ph3-C <sub>ortho'</sub>	-166.278

<sup>b</sup> Fan, C. et al., *J. Chem. Mater.* 2015, 27, 5131–5140.

**Table S2** - Distances (in Å) of C<sub>ortho</sub>, C<sub>ortho'</sub>, C<sub>meta</sub>, C<sub>meta'</sub> from the O atom in TPPO (conformer a) and PPT. Note that, in the case of PPT, the reported distances are related to the molecular moiety considered in spectral calculations (see molecular sketch in Figure S4).

	C <sub>i</sub> -O distances					
	Ph1		Ph2		Ph3	
	TPPO	PPT	TPPO	PPT	TPPO	PPT
<b>C<sub>ortho</sub></b>	3.01	3.02	3.02	3.02	3.02	3.06
<b>C<sub>ortho'</sub></b>	4.01	4.02	4.01	4.03	4.01	4.00
<b>C<sub>meta</sub></b>	4.38	4.40	4.39	4.39	4.39	4.44
<b>C<sub>meta'</sub></b>	5.12	5.13	5.12	5.14	5.12	5.13



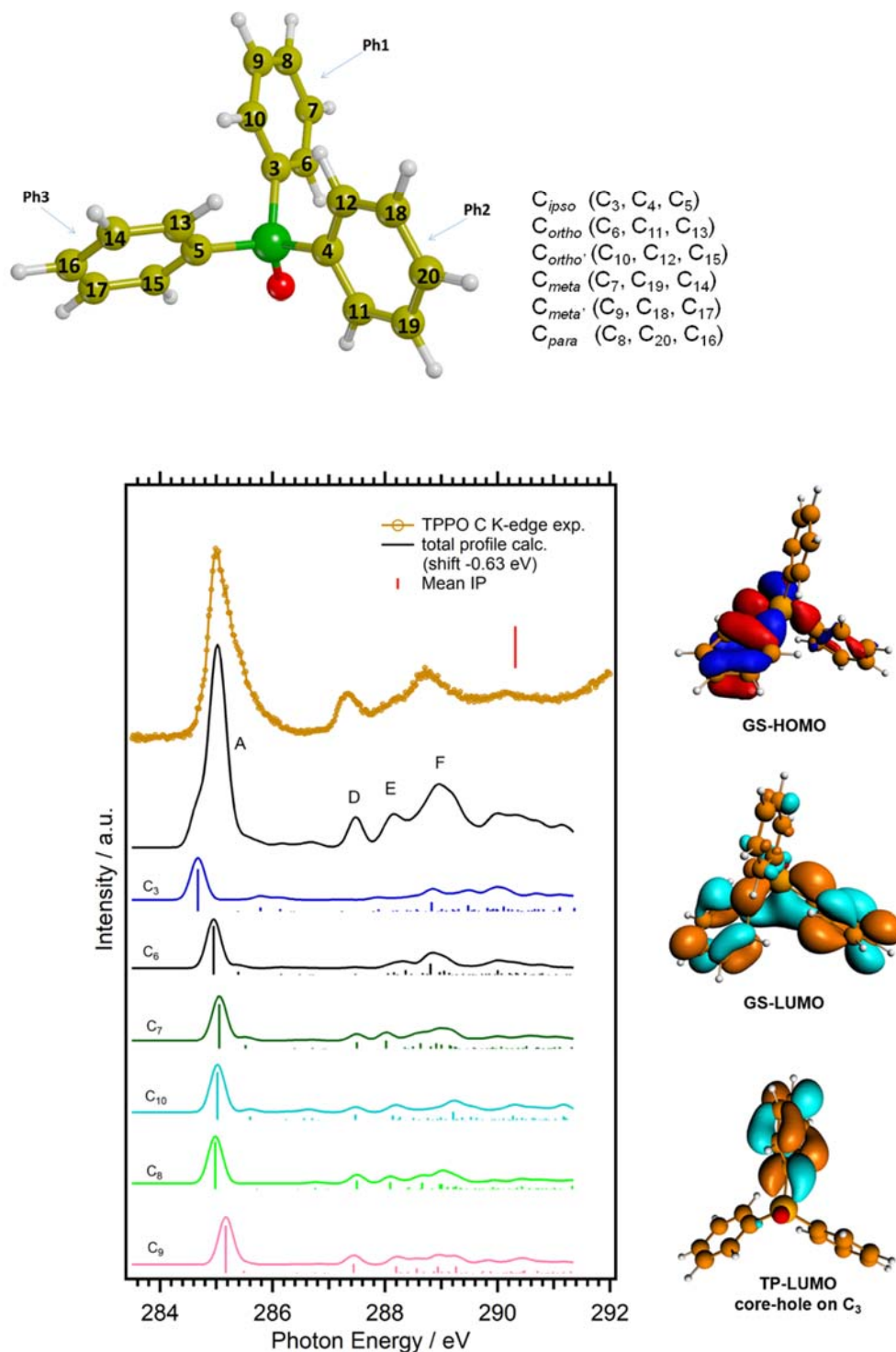
**Figure S2** - Total C K-edge NEXAFS spectra of the two investigated conformers of TPPO: experimental spectrum (gold circles) is compared with the total theoretical line-shapes (solid colored lines). Calculated lines are convoluted by Gaussian profiles with FWHM=0.3 eV. The mean  $\Delta$ KS C 1s ionization thresholds are also shown (colored vertical dashed bars). The calculated peaks have been shifted by  $-0.6$  eV in order to match the first experimental peak.

**Table S3** – Bader charges (in units of the elementary charge,  $e$ ) for the C, P, O atoms of TPPO (a) and also S atom of PPT (b), calculated with the ground-state electron configuration by using the PW86x xc potential.

a) TPPO	Bader charges		
	Ph1	Ph2	Ph3
$C_{\text{ipso}}$	-0.4822	-0.4843	-0.4809
$C_{\text{ortho}}$	0.0069	0.0076	0.0063
$C_{\text{ortho}'}$	0.0022	0.0007	0.0008
$C_{\text{meta}}$	0.0003	-0.0035	-0.0043
$C_{\text{meta}'}$	-0.0002	-0.0021	0.0004
$C_{\text{para}}$	0.0012	0.0052	0.0002
<b>P</b>	2.6063		
<b>O</b>	-1.4392		
b) PPT	Bader charges		
	Ph1	Ph2	Ph3
$C_{\text{ipso}}$	-0.5401	-0.5282	-0.5387

<b>C<sub>ortho</sub></b>	-0.0109	-0.0051	-0.0261
<b>C<sub>ortho'</sub></b>	-0.0262	-0.0217	-0.0153
<b>C<sub>meta</sub></b>	-0.0112	-0.0142	0.0067
<b>C<sub>meta'</sub></b>	-0.0094	0.0026	0.0168
<b>C<sub>para</sub></b>	0.0044	-0.0049	-0.1191
<b>P1</b>	2.7769		
<b>P2</b>	2.7816		
<b>O1</b>	-1.4346		
<b>O2</b>	-1.4324		
<b>S</b>	0.1997		

## C K-edge NEXAFS of TPPO



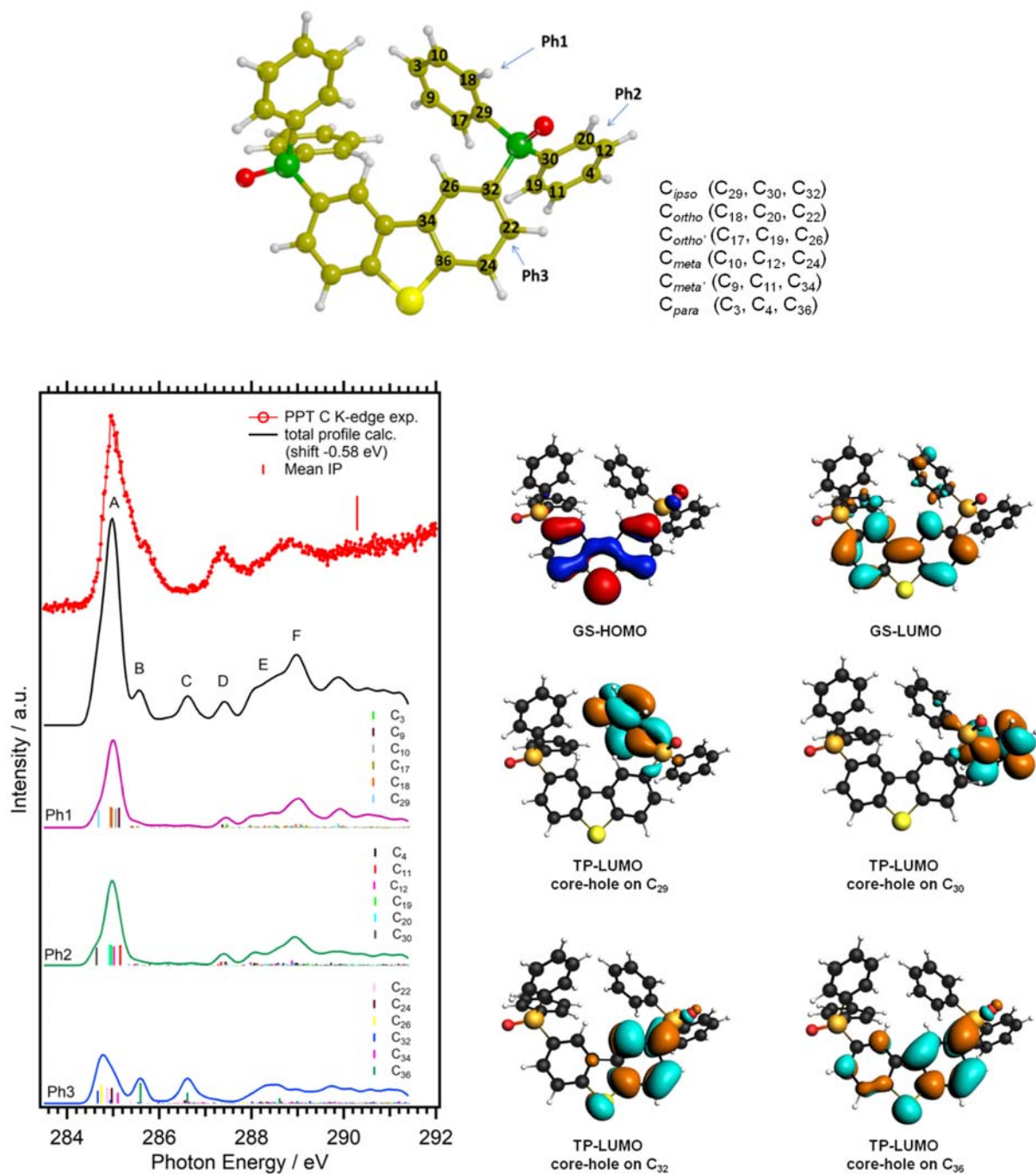
**Figure S3** – C K-edge NEXAFS spectrum of TPPO: experimental spectrum (gold dots) is compared with the total theoretical total line shape (solid black line). Lower panels: spectra of the  $C_i$  contributions (colored solid lines) of each Carbon site. Calculated lines are convoluted by Gaussian profiles with FWHM=0.3 eV. The  $\Delta$ KS C 1s mean ionization threshold is also shown (red vertical bar). The calculated peaks have been shifted by  $-0.63$  eV in order to match the first experimental peak. Right side: plots of ground state (GS) HOMO, LUMO and TP-LUMO with core-hole on  $C_3$ .

**Table S4** – Peak assignments for the C K-edge NEXAFS spectrum of TPPO.

		Ph1		Ph2		Ph3		Assignments	Exp. (eV)
Peak	Site	E(eV) shifted	fx10 <sup>2</sup> <sup>a</sup>	E(eV) shifted	fx10 <sup>2</sup> <sup>a</sup>	E(eV) shifted	fx10 <sup>2</sup> <sup>a</sup>		
A	C <sub>ipso</sub>	284.67	2.64	284.67	2.64	284.67	2.64	LUMO: 1π*(C=C)	284.96
	C <sub>ortho</sub>	284.95	3.07	284.95	3.07	284.96	3.07		
	C <sub>para</sub>	284.98	2.95	284.98	2.95	284.99	2.94		
	C <sub>ortho'</sub>	285.02	2.97	285.02	2.97	285.02	2.97		
	C <sub>meta</sub>	285.05	2.79	285.05	2.80	285.06	2.79		
	C <sub>meta'</sub>	285.17	2.96	285.17	2.97	285.17	2.97		
D	C <sub>meta'</sub>	287.44	0.58	287.44	0.58	287.44	0.58	σ*(C-H)-Rydberg mixed	287.30
	C <sub>ortho'</sub>	287.48	0.33	287.47	0.33	287.47	0.33		
	C <sub>meta</sub>	287.5	0.42	287.5	0.42	287.5	0.42		
	C <sub>para</sub>	287.51	0.55	287.51	0.55	287.51	0.55		
E	C <sub>ipso</sub>	287.89	0.12	287.89	0.12	287.88	0.12	σ*(C-H)-Rydberg mixed	288.24
	C <sub>meta</sub>	288.02	0.51	288.02	0.52	288.02	0.51		
	C <sub>para</sub>	288.09	0.45	288.1	0.45	288.1	0.45		
	C <sub>ortho'</sub>	288.14	0.27	288.14	0.27	288.14	0.27		
	C <sub>meta'</sub>	288.2	0.42	288.2	0.42	288.2	0.42		
F	C <sub>ortho</sub>	288.81	0.70	288.81	0.69	288.81	0.70	π*(C=C)-Rydberg mixed	288.71
	C <sub>ipso</sub>	288.83	0.58	288.83	0.59	288.83	0.58		
	C <sub>meta</sub>	288.91	0.35	288.92	0.37	288.92	0.36		
	C <sub>meta'</sub>	288.94	0.41	288.94	0.41	288.94	0.41		
	C <sub>para</sub>	289	0.33	288.99	0.44	288.99	0.34		
	C <sub>ortho'</sub>	289.22	0.50	289.22	0.50	289.21	0.50		

<sup>a</sup> Only transition with fx10<sup>2</sup> ≥ 1.00x10<sup>-2</sup> have been reported.

## C K-edge NEXAFS of PPT



**Figure S4** – Upper panel: C K-edge NEXAFS spectrum of PPT: experimental spectrum (red circles) is compared with the total theoretical line shape (solid black line). Lower panels: partial  $C_i$  contributions (as colored solid lines) of each phenyl ring (Ph1, Ph2 and Ph3). Calculated lines are convoluted by Gaussian profiles with FWHM=0.4 eV. The  $\Delta$ KS C 1s ionization thresholds are also shown (colored vertical bars). The calculated peaks have been shifted by  $-0.58$  eV to match the first experimental peak. Right side: plots of GS-HOMO, LUMO and selected TP MOs together with the location of the core-hole.

**Table S5** – Peak assignments for the C 1s NEXAFS spectrum of PPT.

PPT		Ph1		Ph2		Ph3		Assignments	Exp. (eV)
Peak	Site	E (eV) shifted	fx10 <sup>2a</sup>	E(eV) shifted	fx10 <sup>2a</sup>	E (eV) shifted	fx10 <sup>2a</sup>		
A	C <sub>ipso</sub>	284.69	5.17	284.64	5.21	284.67	3.68	LUMO: 1π*(C=C)	284.95
	C <sub>ortho</sub>	284.95	5.99	284.92	6.09	284.79	4.52		
	C <sub>para</sub>	284.96	5.82	284.95	5.84	-	-		
	C <sub>ortho'</sub>	284.97	5.86	284.96	5.89	284.75	5.31		
	C <sub>meta</sub>	285.06	5.46	285.02	5.50	284.97	4.46		
	C <sub>meta'</sub>	285.13	5.76	285.15	5.91	285.1	3.07		
	C <sub>ipso</sub>	285.38	7.67x10 <sup>-2</sup>	285.35	3.91x10 <sup>-2</sup>	284.94	1.05	LUMO+1: 2π*(C=C)	-
B	C <sub>para</sub>	-	-	-	-	285.59	5.96	LUMO: 1π*(C=C)	285.69
C	C <sub>meta</sub>	-	-	-	-	286.55	0.93	π*(C=C)	286.67
	C <sub>meta'</sub>	-	-	-	-	286.59	1.05		
	C <sub>para</sub>	-	-	-	-	286.61	3.18	σ*(C-S)	
	C <sub>ortho</sub>	-	-	-	-	286.7	0.40	π*(C=C)	
	C <sub>ipso</sub>	-	-	-	-	286.72	0.16		
	C <sub>ortho'</sub>	-	-	-	-	286.76	0.11	σ*(C-S)	
D	C <sub>ipso</sub>	287.12	3.10x10 <sup>-2</sup>	287.14	1.14x10 <sup>-2</sup>	-	-	σ*(C-S)	287.34
	C <sub>meta'</sub>	287.37	0.89	287.34	1.03	-	-	σ*(C-H)	
	C <sub>ortho'</sub>	287.42	0.18	287.43	0.35	-	-	σ*(C-S)	
	C <sub>ortho</sub>	287.44	7.59x10 <sup>-2</sup>	287.46	5.86x10 <sup>-2</sup>	-	-		
	C <sub>para</sub>	287.46	0.99	287.44	1.05	-	-	σ*(C-H)	
	C <sub>meta</sub>	287.49	0.72	287.49	0.72	-	-		
E	C <sub>ipso</sub>	288.21	0.27	288.2	0.24	288.06	0.13	σ*(C-H)-Rydberg mixed σ*(C-S)	288.22
	C <sub>meta</sub>	288.25	0.18	288.22	9.58x10 <sup>-2</sup>	288.2	0.54		
	C <sub>para</sub>	288.27	9.16x10 <sup>-2</sup>	288.23	0.12	288.61	1.54		
	C <sub>ortho</sub>	288.29	0.55	288.3	0.45	288.17	0.35		
	C <sub>meta'</sub>	288.38	0.53	288.39	0.41	288.43	0.26		
	C <sub>ortho'</sub>	288.39	0.52	288.4	0.40	288.29	0.32		
F	C <sub>ipso</sub>	288.79	0.33	288.69	0.70	288.37	0.37	π*(C=C)-Rydberg mixed	288.64
	C <sub>meta</sub>	288.89	0.61	288.88	1.43	288.91	0.10		
	C <sub>para</sub>	288.89	0.67	288.9	0.55	289.2	0.24	Rydberg	



	<b>C<sub>ortho</sub></b>	288.96	1.02	288.94	0.61	288.71	0.18	$\pi^*(C=C)$ -Rydberg mixed	
	<b>C<sub>meta'</sub></b>	289.07	0.80	289.06	0.73	289.1	0.26	$\pi^*(C=C)$ -Rydberg mixed	
	<b>C<sub>ortho'</sub></b>	289.2	0.61	289.22	0.54	289.02	0.17		
<sup>a</sup> Only transition with $f \times 10^2 \geq 1.00 \times 10^{-2}$ have been reported.									

# S 2p and P 2p Core Level Spectroscopy of PPT Ambipolar Material and Its Building Block Moieties

E. Bernes, G. Fronzoni, M. Stener, A. Guarnaccio,\* T. Zhang, C. Grazioli, F. O. L. Johansson, M. Coreno, M. de Simone, C. Puglia, and D. Toffoli\*



Cite This: *J. Phys. Chem. C* 2020, 124, 14510–14520



Read Online

ACCESS |



Metrics & More

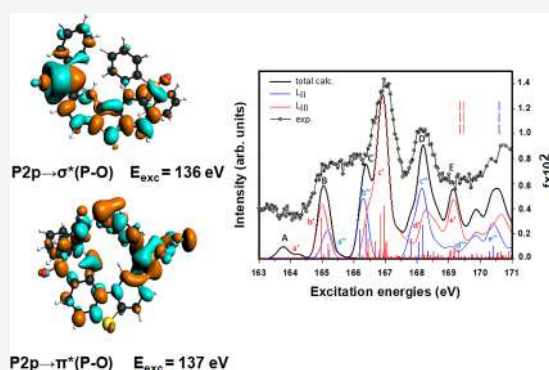


Article Recommendations



Supporting Information

**ABSTRACT:** The near-edge X-ray absorption fine structure (NEXAFS) and X-ray photoelectron (XP) spectra of gas-phase 2,8-bis-(diphenylphosphoryl)dibenzo[*b,d*]thiophene (PPT) and triphenylphosphine oxide (TPPO) have been measured at the S and P  $L_{II,III}$ -edge regions. The time-dependent density functional theory (TDDFT) based on the relativistic two-component zeroth-order regular approximation approach has been used to provide an assignment of the experimental spectra, giving the contribution of the spin–orbit splitting and of the molecular-field splitting to the sulfur and phosphorus binding energies. Computed XP and NEXAFS spectra agree well with the experimental measurements. In going from dibenzothiophene and TPPO to PPT, the nature of the most intense S 2p and P 2p NEXAFS features are preserved; this trend suggests that the electronic and geometric behaviors of the S and P atoms in the two building block moieties are conserved in the more complex system of PPT. This work enables us to shed some light onto the structure of the P–O bond, a still highly debated topic in the chemical literature. Since the S 2p and P 2p NEXAFS intensities provide specific information on the higher-lying localized  $\sigma^*(C-S)$  and  $\sigma^*(P-O)$  virtual MOs, we have concluded that P 3d AOs are not involved in the formation of the P–O bond. Moreover, the results support the mechanism of negative hyperconjugation, by showing that transitions toward  $\sigma^*(P-O)$  states occur at lower energies with respect to those toward  $\pi^*(P-O)$  states.



## 1. INTRODUCTION

During the past decade, numerous advances in the field of organic semiconductor devices have occurred. Currently, a large part of the research is addressing organic light-emitting diodes (OLEDs), and this is due to the synthetic versatility of organic materials, which can be designed with tuned properties, including emission energy, charge transport, and morphological stability.<sup>1,2</sup> In particular, significant efforts have been made on solid-state lighting applications involving phosphorescent OLEDs (PhOLEDs) because of their potential applications in full-color flat-panel displays.<sup>3</sup>

A promising class of host materials in blue PhOLEDs is represented by derivatives of dibenzothiophene (DBT) substituted with diphenylphosphine oxide, such as the 2,8-bis-(diphenylphosphoryl)-dibenzo[*b,d*]thiophene (PPT), an ambipolar phosphorescent electron-transporting material, with sky-blue emission, high emission efficiency, and suitable energy levels.<sup>4–6</sup> The PPT chemical structure (Figure 1, panel c) consists of an electron-rich DBT core functionalized by two electron-withdrawing phosphine oxide groups. It has been seen that the device's efficiency is increased when DBT replaces other  $\pi$ -conjugated core analogues such as carbazole and dibenzofuran.<sup>7</sup> In particular, the polycyclic aromatic DBT building block in PPT promotes the hole transport and reduces

the HOMO–LUMO gap, facilitating the electron injection process and introducing a hole-blocking function in the resulting OLED.<sup>7</sup>

Despite these recent developments toward applications, a detailed understanding of the complex electronic processes involved in the mechanisms underlying their potential device applications is still lacking. To fill this gap, advantage can be taken from a detailed knowledge of their electronic structure, as that provided by the use of core–electron spectroscopies such as XPS (X-ray photoelectron spectroscopy) and NEXAFS (near-edge X-ray absorption fine structure). The latter can be used to probe the electronic structure of complex systems, essentially due to the localized nature of the core-hole which is formed upon core excitation.<sup>8</sup>

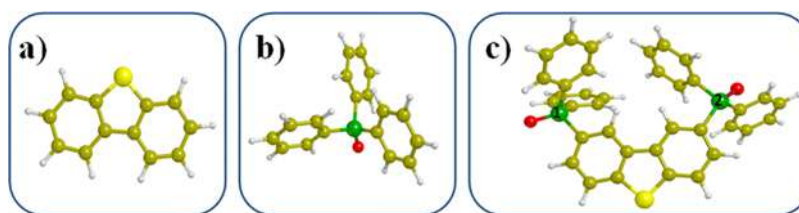
In this respect, we recently studied the C 1s and O 1s edges of the TPPO and PPT molecules;<sup>9</sup> the present paper aims to complete this study by considering the XP and NEXAFS

Received: May 4, 2020

Revised: June 15, 2020

Published: June 15, 2020





**Figure 1.** Chemical structures of (a) dibenzothiophene (DBT), (b) triphenylphosphine oxide (TPPO), and (c) 2,8-bis(diphenylphosphoryl)-dibenzo[*b,d*]thiophene (PPT). O atoms in red, S atoms in yellow, and P atoms in green. The two nonequivalent P atoms of PPT are labeled.

spectra of these two systems at the S and P  $L_{II,III}$ -edges. The comparison with the features of the DBT molecule (see Figure 1, panel a, for the chemical structure) at the S 2p edge<sup>10</sup> will be also considered. To our knowledge, this work represents the first joint experimental and theoretical study of S 2p and P 2p edges on PPT and its DBT and TPPO building blocks. Since the assignment of the NEXAFS P  $L_{II,III}$ -edge absorption features provides specific information on the involvement of P *ns*, *nd* atomic orbitals (AOs) to the virtual molecular states, another outcome of this work is a detailed analysis of the nature of the P–O bond, whose description in terms of atomic orbitals (AOs), contributions of P, O atoms, and the organic residues is still controversial.<sup>11</sup> In our previous study,<sup>9</sup> the O 1s NEXAFS spectra calculations pointed toward a greater participation of O 2p AOs to occupied molecular orbitals (MOs) and supported the mechanism of negative hyperconjugation as well as the stabilization of the  $\pi$  symmetry interaction in the P–O bond with respect to the  $\sigma$  one in the occupied MOs. We expect that the present results on the P  $L_{II,III}$ -edge can support these findings.

Both the coupling between different excitation channels arising from the 2p degenerate core-holes and relativistic spin–orbit (SO) coupling effects need to be included in the computational protocol used to simulate the NEXAFS spectra at the  $L_{II,III}$ -edge region, to obtain a quantitative description of the series of transitions converging to the  $L_{II}$  and  $L_{III}$ -edges. In this paper we used TDDFT and the relativistic two-component zeroth-order regular approximation approach (ZORA), as implemented in the ADF (Amsterdam density functional) code for the treatment of core electron excitations. Core ionization energies (IEs), measured by XPS, are also useful for material characterization due to their sensitivity to the local chemical and physical environment of the ionized atomic site. In the ZORA-TDDFT scheme, the absolute IP's values are identified as minus the ground-state Kohn–Sham spinors eigenvalues, while the intensity of the primary lines has been set arbitrary to unity multiplied by the number of symmetry equivalent sites (for the two P atoms in PPT). This representation stems from the difficulty to calculate the  $\mu_{fi}$  dipole transition moment between the core orbital and the continuum orbital of the unbound photoelectron, when conventional basis sets, like Gaussian or Slater functions, are used, and is supported by the sufficiently high photon energy employed in the S 2p and P 2p XP experiments (see the Experimental Section) with respect to the S 2p and P 2p binding energies (BEs).

The plan of this paper is the following: Sections 2 and 3 describe the experimental setup and the computational method employed, respectively. A discussion of the results is presented in Section 4, where the assignment of S 2p and P 2p XP and NEXAFS spectra of TPPO and PPT is carried out, while

conclusions and perspectives are summarized in the final Section 5.

## 2. EXPERIMENTAL SECTION

The investigated molecules were purchased from Sigma-Aldrich (TPPO, with purity 98%) and from Lumtec (PPT, with purity >99%). At room temperature both TPPO and PPT are crystalline solids, and the gas phase was reached by sublimation in vacuum employing a in-house built resistive furnace. The sublimation temperatures were 128 °C for TPPO and 251 °C for PPT. Both temperatures were reached after a gradient purification treatment of at least 20 h.

The experiments were carried out at the Gas Phase beamline of the Elettra-Sincrotrone Trieste<sup>12</sup> by using an end-station equipped with a Scienta SES-200 electron analyzer.<sup>13</sup> The angle between the detection direction and the electric vector of the linearly polarized light was 54° (magic angle configuration). Photoemission spectra of the S 2p core levels were collected using a photon energy of 260 eV and an energy resolution of about 111 meV; S 2p spectra of SF<sub>6</sub> were also collected for energy calibration (180.21 eV for 2p<sub>3/2</sub> and 181.5 eV for 2p<sub>1/2</sub>).<sup>14,15</sup> P 2p spectra were collected at a photon energy of 220 eV with an energy resolution of about 114 meV; the energy calibration was performed by measuring in the same experimental window both S 2p and P 2p doublets and calibrating with respect to the previously calibrated S 2p XPS doublet of the corresponding molecule.

NEXAFS spectra at the S, P  $L_{II,III}$ -edges were acquired in the total ion yield (TIY) mode with an electron multiplier and normalized with respect to the photon flux by using a calibrated Si photodiode (AxVU100 IRD). In the case of S, the calibration of the photon energy was performed with respect to the characteristic  $T_{1u}(a_{1g})-3/2 T_{1u}(a_{1g})-1/2$  transition of SF<sub>6</sub> (respectively at 172.5 and 173.6 eV<sup>16</sup>) collected simultaneously. The photon energy resolution was 80 meV. The P  $L_{II,III}$ -edge was calibrated by taking simultaneous spectra of the samples and of Xe, whose characteristic Xe 4p transition (141.8 eV) is well-known.<sup>17</sup> The peak resolution of the P  $L_{II,III}$ -edge spectra was equal to 50 meV for both TPPO and PPT.

The experimental XPS data were analyzed by means of the XPST (X-ray photoelectron spectroscopy tools) package for IGOR PRO by Dr. Martin Schmid. The molecular-field splitting was extracted from the S 2p and P 2p spectra by applying an empirical model similar to the one by Giertz et al.<sup>18</sup> In the first step, the data were fitted using one single asymmetric Voigt function for S 2p<sub>1/2</sub> and P 2p<sub>1/2</sub> peaks. The same asymmetric Voigt function was then superimposed twice on the S 2p<sub>3/2</sub> and P 2p<sub>3/2</sub> peaks. Then appropriate shifts have been applied to the calculated eigenvalues in order to better align them with the corresponding experimental fitted curves.

### 3. COMPUTATIONAL DETAILS

The geometry optimization of TPPO and PPT has been carried out at the density functional theory (DFT<sup>19</sup>) level, within the local-density approximation (LDA) by using the Vosko, Wilk, Nusair (VWN) exchange and correlation (xc) functional<sup>20</sup> and the triple  $\zeta$  polarized (TZP) basis set of Slater type orbitals (STOs) from the ADF (Amsterdam density functional) database.<sup>21,22</sup> No symmetry constraints were imposed during the optimization. Indeed, as revealed by crystallographic studies, TPPO has no symmetry because the rotations of the phenyl rings about the P–C bonds are not identical.<sup>23,24</sup> Moreover, the O–P–C angles are greater than the tetrahedral angle due to the partial multiple bond character of the P–O bond, while the C–P–C angles are smaller than it.<sup>24</sup> Analogous considerations can be made in the case of PPT, where the lack of symmetry is a consequence of steric avoidance between the two diphenylphosphine oxide moieties. Two different conformations have been taken into account for TPPO, one derived from X-ray diffraction data<sup>25</sup> and one predicted in vacuo at the MP2/cc-pVTZ level.<sup>26</sup> Since we found out that the C 1s NEXAFS spectra of the two conformers are quantitatively very similar,<sup>9</sup> we chose to perform the computation and the assignment of the P 2p NEXAFS spectrum of TPPO only on the structure derived from the XRD data to which also all results of our previous study refer. The optimized Cartesian coordinates of DBT, TPPO, and PPT are reported in the [Supporting Information](#).

The S 2p and P 2p NEXAFS spectra calculations have been performed within TDDFT by using the two-component ZORA approximation<sup>27</sup> implemented in the ADF code<sup>21,22,28</sup> within the adiabatic local density approximation (ALDA) to the exchange-correlation (xc) kernel. Here, we will only recall the main steps of the TDDFT formalism when dealing with core electron excitations, referring the reader to the literature<sup>29,30</sup> for a more detailed account of the method.

Within linear response TDDFT,<sup>31</sup> the excitation spectrum is obtained through the solution of the following eigenvalue equation by means of Davidson's iterative algorithm:<sup>32</sup>

$$\Omega F_1 = \omega_1^2 F_1 \quad (1)$$

The elements of the  $\Omega$  matrix are given by

$$\Omega_{i\sigma r, b j \tau} = \delta_{\sigma r} \delta_{ij} \delta_{ab} (\epsilon_a - \epsilon_i)^2 + 2\sqrt{(\epsilon_a - \epsilon_i)} \frac{\partial F_{ia}}{\partial P_{jb}} \sqrt{(\epsilon_b - \epsilon_j)} \quad (2)$$

where indices  $i$  and  $j$  denote occupied spinors in the KS ground state, while indices  $a$  and  $b$  denote virtual spinors and  $\epsilon_i$  and  $\epsilon_a$  are the corresponding KS energies. The Fock and the density matrix are denoted by  $F$  and  $P$  while  $\frac{\partial F_{ia}}{\partial P_{jb}}$  are the elements of the coupling matrix. Squared excitation energies  $\omega_1^2$  are then obtained as the eigenvalues of eq 1, while oscillator strengths can be extracted from the eigenvectors  $F_1$ . However, due to its limitation to the extraction of a relatively small number of lowest eigenvalues and eigenvectors, Davidson's algorithm is not able to efficiently solve eq 1, as core excitation energies lie very high in the excitation spectrum. This limitation can be overcome by invoking the core–valence separation approximation (CVS) introduced by Cederbaum et al.,<sup>33,34</sup> the 1h-1p space spanned by the solutions of eq 1 can be reduced by including only the most important configurations.<sup>35</sup> Therefore, within this approximation, indices  $i$  and  $j$  run only over the

core–shell under study (in the case of 2p excitations, over the two subshells, namely 2p<sub>1/2</sub> and 2p<sub>3/2</sub>, allowing a mixing of only those initial states). As a consequence, eq 1 is solved in a smaller 1h-1p subspace, and the core excitations now correspond to its lowest roots, hence allowing an effective use of Davidson's algorithm. This approximation neglects the coupling between core excitations and valence excitations to the continuum, and the treatment of the continuum states does not need to be included in the computational scheme. In addition, one can expect that the coupling between core excitations and valence ionizations is very small.<sup>35</sup>

For the nonexcited atoms, an all-electron DZP basis set, optimized for ZORA calculations<sup>36</sup> and available from the ADF database, has been used. For both P and S excited atoms, two shells of s, p, d, and f diffuse functions with exponents obtained with the even-tempered criterion ( $\beta = 1.7$ ) have been added to the QZ4P ZORA basis set to obtain a [13s, 9p, 5d, 6f] basis, referred to as ET-QZ4P-2DIFFUSE. We include diffuse functions in the basis set to describe transitions to diffuse orbitals (Rydberg excited states) close to the ionization threshold. Two different xc functionals have been tested in the self-consistent field (SCF) calculations: the LB94 xc potential<sup>37</sup> with the ground-state electron configuration and the hybrid B3LYP.<sup>38–40</sup> Since the former gives a slightly better agreement with the experimental data for both XP and NEXAFS spectra, in the following section, we will discuss the LB94 results and provide the B3LYP outcomes in the [Supporting Information](#). Furthermore, the assignment of the experimental features of the NEXAFS spectra to 1h-1p core-excited states is not significantly affected by the actual choice of the xc potential. A note on the computational efficiency of the CVS-TDDFT method when applied to large molecular systems, such as PPT, is in order. In such cases, the density of the manifold of core-excited states is such that the number of roots of eq 1 that need to be extracted to cover the excitation energy range up to the P L<sub>II/III</sub> ionization threshold is very high (>1000) so that even within the CVS approximation the Davidson's iterative algorithm could become numerically unstable. In this work, the computed P 2p TDDFT NEXAFS profiles are therefore limited to the energy interval determined by the largest roots that can be safely computed iteratively (700 and 1000 for TPPO and PPT, respectively). This is not a problem when employing the LB94 potential but, due to the specific implementation within ADF, poses several limitations when using hybrid functionals such as B3LYP.

For both xc functionals, the calculated S 2p and P 2p NEXAFS spectra have been deconvoluted into the two manifolds of excited states converging to the L<sub>III</sub> and L<sub>II</sub> edges. This has been done by multiplying the oscillator strength of each transition by the coefficient squared of the configurations describing excitations from the 2p<sub>3/2</sub> and 2p<sub>1/2</sub> levels. In the calculation of the P 2p NEXAFS spectra, we included the coupling among the excitation channels of both P centers, since preliminary calculations revealed that their coupling is mandatory for achieving a quantitative reproduction of the experimental NEXAFS intensity distribution. Core excitation energies are computed from scaled ZORA spinor eigenvalues.

The raw NEXAFS spectra have been convoluted with Gaussian functions of constant full width at half-maximum (fwhm) values (equal to 0.3 and 0.5 eV for S 2p and P 2p NEXAFS spectra, respectively). The use of Gaussian functions to smooth calculated discrete lines allows for a pragmatic

inclusion of an average lifetime of the core-excited states, beside facilitating the comparison with the experimental data. Although vibrational effects can also affect the line-shape and broadening of the core-ionized and -excited spectra, they have not been considered in the present work.

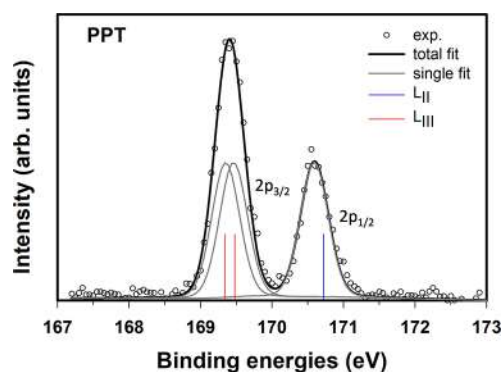
Theoretical ionization threshold values are identified as minus the DFT eigenvalues obtained within the ZORA approximation and the LB94 xc functional

$$IP = -\varepsilon_{2p} \quad (3)$$

where  $\varepsilon_{2p}$  represents the eigenvalue of the  $2p_{1/2,3/2}$  spin orbitals. These IP values are also employed to simulate the S 2p and P 2p XP spectra; when comparing with the experimental measurements, the theoretical IPs have been shifted to lower BEs by 1.23 eV (S 2p PPT), 0.90 eV (P 2p TPPO), and 1.02 eV (P 2p PPT). The energy scale of the calculated NEXAFS spectra has been shifted to align the energy position of the maxima of the first absorption band. Applied shifts were equal to +0.3, +2.25, and +2.20 eV for the S  $L_{II,III}$ -edge excitations in PPT and P  $L_{II,III}$ -edge excitations in TPPO and PPT, respectively.

#### 4. RESULTS AND DISCUSSION

In the following two subsections we will analyze our calculations to assign the experimental XP and NEXAFS



**Figure 2.** S 2p XP spectrum of PPT. Circles: experimental data; black line: total fit obtained according to the procedure described in the text. Vertical colored bars: LB94 IPs shifted by  $-1.23$  eV for a better matching with the experimental curve. The light gray curves are the  $L_{II}$  (centered around the blue vertical bar) and  $L_{III}$  (centered around the red vertical bars) S 2p components obtained by the fitting procedure using a single and two asymmetric Voigt functions, respectively.

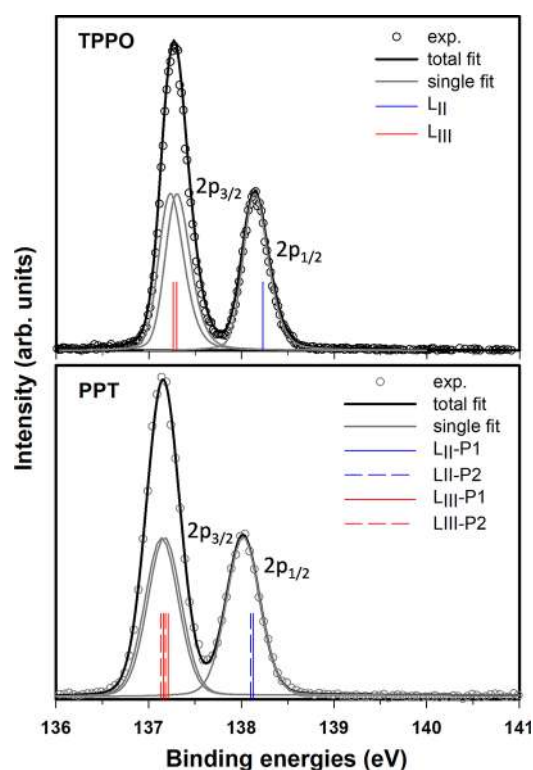
**Table 1.** Theoretical LB94 S 2p IPs and Experimental BEs of PPT<sup>a</sup>

edge	PPT	
	theory <sup>b</sup>	single fit
$L_{II}$	170.72	170.59
$L_{III}$	169.48	169.46
$L_{III}$	169.34	169.35

<sup>a</sup>All values are expressed in eV. <sup>b</sup>Calculated IPs shifted by  $-1.23$  eV.

spectra of TPPO and PPT. We will focus on the XP spectra in the first subsection, while NEXAFS spectra will be considered in the final subsection.

The intensity distribution of the  $L_{II,III}$ -edge spectra is governed by the interplay of two main effects: the relativistic



**Figure 3.** P 2p XP spectra of TPPO (upper panel) and PPT (lower panel). Circles: experimental data; black line: total fit obtained according to the procedure described in the text. Vertical colored bars: LB94 IPs shifted by  $-0.90$  and  $-1.02$  eV for TPPO and PPT, respectively, for a better matching with the experimental curves. The light gray curves are the  $L_{II}$  (centered around the blue vertical bar) and  $L_{III}$  (centered around the red vertical bars) S 2p components obtained by the fitting procedure using a single and two asymmetric Voigt functions, respectively.

**Table 2.** Theoretical LB94 P 2p IPs and Experimental BEs of TPPO and PPT<sup>a</sup>

edge	TPPO		PPT		
	theory <sup>b</sup>	single fit	theory <sup>c</sup>		single fit
			P1	P2	
$L_{II}$	138.23	138.14	138.13	138.10	138.01
$L_{III}$	137.30	137.30	137.21	137.18	137.17
$L_{III}$	137.26	137.23	137.16	137.13	137.14

<sup>a</sup>All values are in eV. <sup>b</sup>Calculated IPs shifted by  $-0.90$  eV. <sup>c</sup>Calculated IPs shifted by  $-1.02$  eV.

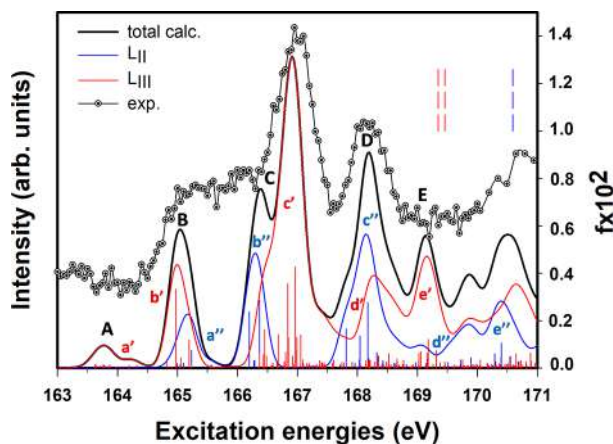
SO interaction and the *molecular-field splitting*. SO coupling lowers the degeneracy of the 2p core hole by splitting the 2p core orbital energies into the two SO components  $2p_{3/2}(L_{III})$  and  $2p_{1/2}(L_{II})$ . The molecular-field splitting further lowers the degeneracy of the  $2p_{3/2}$  state and produces a splitting into two components, as a consequence of the anisotropic molecular potential. Moreover, configuration mixing plays an important role in determining the correct intensity distribution among the 2p transitions. All these effects are included in the computational method employed.

**4.1. Assignment of the Experimental S 2p and P 2p XP Spectra.** S 2p XP spectra of PPT are shown in Figure 2, together with the results of the fitting procedure detailed in Section 2, while a comparison with the S 2p results for DBT<sup>10</sup> has been included in the Supporting Information (Figure S1).

Table 3. Peak Assignments of the S L<sub>II,III</sub>-Edge NEXAFS Spectrum of PPT<sup>a</sup>

peak/ subpeak	calcd				exptl	peak/ subpeak	calcd				exptl
	<i>E</i> -shifted <sup>b</sup> (eV)	<i>f</i> × 10 <sup>2</sup>	EDGE	assignment	<i>E</i> (eV)		<i>E</i> -shifted <sup>b</sup> (eV)	<i>f</i> × 10 <sup>2</sup>	EDGE	assignment	<i>E</i> (eV)
A/a'	163.62	1.39 × 10 <sup>-2</sup>	L <sub>III</sub>	155 A (LUMO+2)/ π*(C=C)	-		166.80	6.78 × 10 <sup>-2</sup>		170 A/σ*(P-O)- Rydberg mixed	
	163.65	1.25 × 10 <sup>-2</sup>		156 A/π*(C=C)			166.84	0.36		172 A/Rydberg	
	163.76	1.41 × 10 <sup>-2</sup>		155 A			166.85	0.13		171 A	
	163.77	1.13 × 10 <sup>-2</sup>		157 A/π*(C=C)			166.92	4.99 × 10 <sup>-2</sup>		172 A	
	163.85	2.90 × 10 <sup>-3</sup>		158 A/π*(C=C)			166.94	1.66 × 10 <sup>-2</sup>		173 A/σ*(C-H)- Rydberg mixed	
	163.87	1.41 × 10 <sup>-2</sup>		157 A			166.97	0.43		172 A	
	163.99	4.05 × 10 <sup>-3</sup>		158 A			166.99	0.14		172 A	
	164.20	1.06 × 10 <sup>-2</sup>		163 A/π*(C=C)			167.06	0.14		174 A/σ*(C-H)- Rydberg mixed	
	164.32	1.08 × 10 <sup>-2</sup>					167.19	2.18 × 10 <sup>-2</sup>			
B/a''	164.99	1.92 × 10 <sup>-2</sup>	L <sub>II</sub>	155 A	165.30	D/c''	167.20	1.18 × 10 <sup>-2</sup>	L <sub>II</sub>	168 A	168.00
	165.02	6.41 × 10 <sup>-3</sup>		156 A			167.81	0.17		171 A	
	165.10	2.33 × 10 <sup>-2</sup>		157 A			168.04	0.14		172 A	
	165.22	3.64 × 10 <sup>-3</sup>		158 A			168.18	0.31			
	165.23	8.98 × 10 <sup>-2</sup>					168.19	6.76 × 10 <sup>-2</sup>			
	165.56	1.21 × 10 <sup>-2</sup>		163 A			168.32	6.93 × 10 <sup>-2</sup>		173 A	
B/b'	164.82	3.58 × 10 <sup>-2</sup>	L <sub>III</sub>	164 A/σ*(C-S)- Rydberg mixed			168.43	3.28 × 10 <sup>-2</sup>		174 A	
	164.97	0.34					168.44	1.35 × 10 <sup>-2</sup>			
	165.06	4.39 × 10 <sup>-2</sup>		165 A/σ*(C-S)- Rydberg mixed		D/d'	168.35	5.88 × 10 <sup>-2</sup>	L <sub>III</sub>	189 A/Rydberg	
	165.19	0.22		164 A			168.52	6.31 × 10 <sup>-2</sup>		191 A/σ*(C-H)- Rydberg mixed	
C/b''	166.20	0.24	L <sub>II</sub>	164 A	-	E/e'	169.03	6.39 × 10 <sup>-2</sup>	L <sub>III</sub>	196 A/σ*(C-H)- Rydberg mixed	-
	166.27	3.03 × 10 <sup>-2</sup>					169.06	8.45 × 10 <sup>-2</sup>		197 A/σ*(C-H)- Rydberg mixed	
	166.29	3.27 × 10 <sup>-2</sup>		165 A			169.16	6.63 × 10 <sup>-2</sup>		198 A/σ*(C-H)- Rydberg mixed	
	166.37	0.36					169.18	0.12		200 A/σ*(C-H)- Rydberg mixed	
C/c'	166.44	5.80 × 10 <sup>-2</sup>	L <sub>III</sub>	168 A/σ*(P-O)- Rydberg mixed	166.75		169.32	7.27 × 10 <sup>-2</sup>			
	166.45	0.18									
	166.48	5.17 × 10 <sup>-2</sup>		167 A/σ*(C-H)- Rydberg mixed							
	166.68	0.14		171 A/σ*(C-H)- Rydberg mixed							

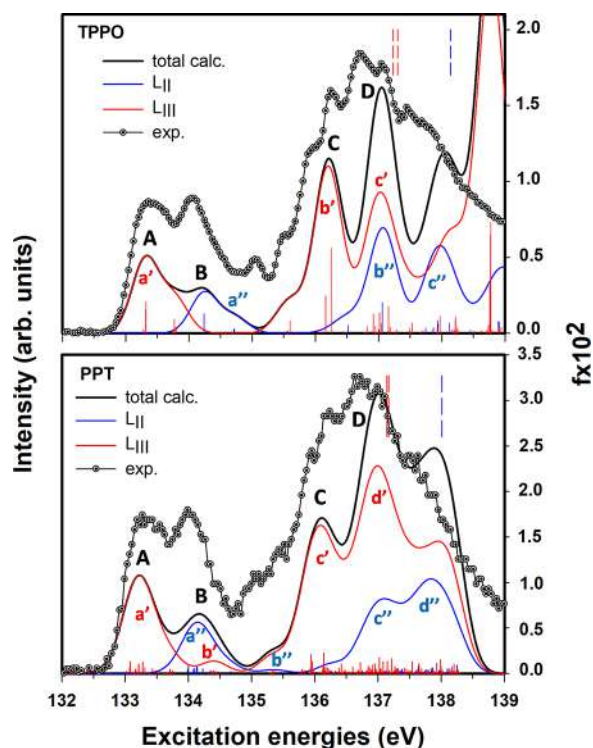
<sup>a</sup>Only the main transitions, computed at the TDDFT LB94 level, are reported. <sup>b</sup>Calculated excitation energies shifted by +0.3 eV to match the energy position of the first experimental peak.



**Figure 4.** S L<sub>II,III</sub>-edge NEXAFS spectrum of PPT. Circles: experimental data; black solid line: LB94 TDDFT results. Deconvolution of the calculated S 2p spectrum into excitations converging to the L<sub>III</sub> (red solid line and vertical red bars) and L<sub>II</sub> (blue solid line and vertical blue bars) edges is also shown. The energy scale of the theoretical data has been shifted by +0.3 eV to match the position of the first experimental peak. Blue and red vertical bars denote the experimental L<sub>II,III</sub>-edges. The left vertical axis refers to the experimental intensities (plotted in arbitrary units), while the right vertical axis refers to computed oscillator strengths (multiplied by a factor of 100).

The experimentally fitted BEs together with the LB94 IPs of PPT are also listed in Table 1.

We assign the more intense lower-energy photoelectron peak of Figure 2 to the <sup>2</sup>P<sub>3/2</sub> states, while the peak at 170.6 eV is assigned to the <sup>2</sup>P<sub>1/2</sub> state. The fitting reproduces very well the experimental data and permits the extraction of the molecular-field splitting of the S 2p<sub>3/2</sub> component. The experimental trend is qualitatively reproduced by theory, although the KS eigenvalues overestimate the fitted BEs by about 1.23 eV; this can be ascribed to the too attractive character of the LB94 xc potential. Within standard TDDFT, the 1h-1p excitation space is built from the KS ground-state orbitals (which are therefore not relaxed), and the IPs of the 2p electrons, according to Koopmans' theorem, are identified as minus the KS energies of the relative spinors. The discrepancy with the observed trend could be in principle recovered by ΔKS calculation of IPs<sup>28</sup> as well as by the use of relaxed orbitals.<sup>29</sup> On the other hand, the SO splitting (Δ<sub>3/2,1/2</sub>) value, computed as the difference between the energy of the <sup>2</sup>P<sub>1/2</sub> state and the average energy of the <sup>2</sup>P<sub>3/2</sub> states, is in very good agreement with the experimental value, being the discrepancy of the two sets of data of 120 meV (1.31 eV computed, 1.19 eV experimental), and confirms *a posteriori* the accuracy of the computational method employed. The computed molecular-field splitting, equal to 140 meV, is in good agreement with the experimental value derived from the fit (equal to 110 meV).



**Figure 5.** P  $L_{II,III}$ -edge NEXAFS spectra of TPPO (upper panel) and PPT (lower panel). Circles: experimental data; black solid line: LB94 TDDFT results. Deconvolution of the calculated spectra into excitations converging to the  $L_{III}$  (red solid line and vertical red bars) and  $L_{II}$  (blue solid line and vertical blue bars) edges is also shown. The energy scale of the theoretical data has been shifted by +2.25 and +2.20 eV, respectively, for TPPO and PPT to match the position of the first experimental peak. The experimental  $L_{II,III}$ -edges are also shown (blue and red vertical dashed bars) within each spectrum. The left vertical axis refers to the experimental intensities (plotted in arbitrary units), while the right vertical axis refers to computed oscillator strengths (multiplied by a factor of 100).

From a comparison of the S 2p spectrum of PPT with that of DBT (see Figure S1), one sees that the two spectra are very similar, as well as their SO splitting. A small BE increase (of about 0.1 eV) of the photoelectron peaks is found in going from DBT to PPT (compare Table S1 and Table 1): this reflects a small local change of the electron density around sulfur (lower shielding effect), as a consequence of the addition of two electron-withdrawing phosphine oxide moieties in PPT. The value of the molecular-field splitting predicted by theory is similar for DBT and PPT; this is in line with the comparable geometrical environment of the S atom in the two molecules.

Figure 3 displays the X-ray photoelectron spectra of the P 2p core levels of TPPO and PPT, together with the results of the fit of the experimental line-shapes as described in Section 2. Theoretical LB94 IPs are also reported in the figure, while B3LYP IPs are reported in Table S2 of the Supporting Information.

The experimental BEs are reported in Table 2 along with the computed LB94 P 2p IPs (see Figure 1, panel c, for the labeling of the two nonequivalent P atoms).

An analysis of Figure 3 reveals that the shape of the photoelectron spectra of the two molecules is very similar; indeed, a stronger photoelectron peak, assigned to the  $^2P_{3/2}$  states, is observed at lower BEs, while the band at higher energy is assigned to the  $^2P_{1/2}$  state. Their energy separation

**Table 4.** Peak Assignments of the P  $L_{II,III}$ -Edge NEXAFS Spectrum of TPPO<sup>a</sup>

peak/ subpeak	calcd		EDGE	assignment	exptl E (eV)
	E-shifted <sup>b</sup> (eV)	$f \times 10^2$			
A/a'	133.32	0.21	$L_{III}$	74 A (LUMO)/ $\pi^*(C=C)$	133.16
	133.77	$9.19 \times 10^{-2}$			
	133.79	$4.14 \times 10^{-2}$			
B/a''	134.24	0.13	$L_{II}$	75 A/ $\pi^*(C=C)$	134.08
	134.71	$2.41 \times 10^{-2}$			
	134.72	$2.96 \times 10^{-2}$			
	135.60	$8.44 \times 10^{-2}$			
C/b'	135.60	$8.44 \times 10^{-2}$	$L_{III}$	80 A/ $\sigma^*(C-H)$ - Rydberg mixed	135.48
	136.16	0.25			
	136.25	0.58			
	136.52	$5.12 \times 10^{-2}$			
	136.25	0.58			
D/b''	136.52	$5.12 \times 10^{-2}$	$L_{II}$	80 A	136.68
	137.06	0.21			
D/c'	136.82	$4.33 \times 10^{-2}$	$L_{III}$	85 A/ $\sigma^*(C-H)$ - Rydberg mixed	137.48
	136.92	0.13			
	136.95	$4.92 \times 10^{-2}$			
	137.02	0.17			
	137.05	$9.25 \times 10^{-2}$			
	137.16	0.33			
	137.18	$5.46 \times 10^{-2}$			
	137.20	0.17			
	137.05	$9.25 \times 10^{-2}$			
	137.16	0.33			
	137.18	$5.46 \times 10^{-2}$			

<sup>a</sup>Only the main transitions, computed at the TDDFT LB94 level, are reported. <sup>b</sup>Calculated excitation energies shifted by +2.25 eV to match the energy position of the first experimental peak.

corresponds to the SO splitting of the 2p SO components. A small decrease (of about 100 meV) of the  $L_{III}$  and  $L_{II}$  experimental BEs is found in going from TPPO to PPT (see Table 2), as a result of the replacement of a single phenyl ring of TPPO with one condensed ring of the DBT moiety in PPT. The experimental data are well reproduced by the fits, which permit the obtaining of an experimental estimate of the molecular-field splitting of the P  $2p_{3/2}$  states. Moreover, the fact that the two diphenylphosphine oxide moieties point out toward different directions due to steric hindrance causes a splitting of the P  $2p_{3/2}$  signal into two very narrowly spaced components, separated by 30 meV. Furthermore, Table 2 reveals that the shielding effect on the BEs is lower than the molecular-field splitting. We can conclude that the energy separation of the two doublets of the lower-energy band of the PPT XP spectrum is mainly determined by the anisotropic molecular potential which splits the P  $2p_{3/2}$  states by about 50 meV.

As concerns the comparison between experiment and theory, the discrepancy between the theoretical IPs and the fitted BEs is about 0.9 eV for TPPO and about 1 eV for PPT; the reason for these discrepancies in the absolute values is again due to the too attractive character of the LB94 xc potential, and considerations analogous to those made when discussing the S 2p XP spectra can be made here. However, the computed SO splitting among the  $L_{III}$  and  $L_{II}$  P 2p thresholds is in good agreement with the experimental value. Indeed, the computed SO splitting is equal to 950 meV for both molecules, in line with the experimental value of about 870 meV; this still confirms that relativistic effects are rather well accounted for by the ZORA formalism. The calculated molecular-field splitting

Table 5. Peak Assignments of the P L<sub>II,III</sub>-Edge NEXAFS Spectrum of PPT<sup>a</sup>

peak/subpeak	calcd			EDGE	assignment	exptl				
	core-hole site	<i>E</i> -shifted <sup>b</sup> (eV)	<i>f</i> × 10 <sup>2</sup>			<i>E</i> (eV)				
A/a'	P2	133.07	6.40 × 10 <sup>-2</sup>	L <sub>III</sub>	154 A (LUMO+1)/π*(C=C)	133.39				
	P2	133.08	0.13							
	P1	133.16	3.61 × 10 <sup>-2</sup>							
	P1	133.17	5.03 × 10 <sup>-2</sup>							
	P1	133.21	0.10							
	P1	133.22	9.98 × 10 <sup>-2</sup>							
	P2	133.28	0.13							
	P2	133.31	6.29 × 10 <sup>-2</sup>							
	P2	133.43	5.27 × 10 <sup>-2</sup>							
	P1	133.71	4.76 × 10 <sup>-2</sup>							
	P1	133.75	3.09 × 10 <sup>-2</sup>							
	B/a''	P2	133.99				4.63 × 10 <sup>-2</sup>	L <sub>II</sub>	154 A	133.99
		P2	134.00				8.06 × 10 <sup>-2</sup>			
P1		134.08	4.29 × 10 <sup>-2</sup>							
P1		134.09	5.42 × 10 <sup>-2</sup>							
P1		134.13	6.77 × 10 <sup>-2</sup>							
P1		134.14	6.85 × 10 <sup>-2</sup>							
P2		134.23	4.98 × 10 <sup>-2</sup>							
P2		134.35	3.35 × 10 <sup>-2</sup>							
B/b'		P2	134.38	2.19 × 10 <sup>-2</sup>	L <sub>III</sub>	164 A/π*(C=C)				
		P1	134.40	2.64 × 10 <sup>-2</sup>						
	P2	134.43	2.54 × 10 <sup>-2</sup>							
C/b''	P1	135.32	9.39 × 10 <sup>-3</sup>	L <sub>II</sub>	164 A	135.23				
	P2	135.35	9.23 × 10 <sup>-3</sup>							
C/c'	P1	135.41	3.90 × 10 <sup>-3</sup>	L <sub>III</sub>	165 A/Rydberg	135.91				
	P1	135.33	4.24 × 10 <sup>-2</sup>							
	P2	135.36	4.01 × 10 <sup>-2</sup>							
	P1	135.80	4.61 × 10 <sup>-2</sup>							
	P2	135.93	6.18 × 10 <sup>-2</sup>							
	P1	135.94	0.21							
	P1	135.96	0.14							
	P1	135.97	5.02 × 10 <sup>-2</sup>							
	P1	136.13	9.16 × 10 <sup>-2</sup>							
	P1	136.14	0.23							
	P1	136.15	7.05 × 10 <sup>-2</sup>							
	P2	136.19	6.80 × 10 <sup>-2</sup>							
	P1	136.20	6.15 × 10 <sup>-2</sup>							
	P2	136.23	5.00 × 10 <sup>-2</sup>							
	P1	136.38	5.86 × 10 <sup>-2</sup>							
	D/c''	P2	136.42				0.11	L <sub>II</sub>	173 A/σ*(C-H)-Rydberg mixed	136.72
P1		136.48	3.67 × 10 <sup>-2</sup>							
P2		136.75	3.18 × 10 <sup>-2</sup>							
P2		136.89	6.20 × 10 <sup>-2</sup>							
P1		137.04	5.18 × 10 <sup>-2</sup>							
P1		137.08	0.16							
P1		137.12	3.14 × 10 <sup>-2</sup>							
D/d'		P2	136.60	7.24 × 10 <sup>-2</sup>	L <sub>III</sub>	174 A/Rydberg	136.99			
		P1	136.68	3.94 × 10 <sup>-2</sup>						
		P2	136.72	0.12						
	P2	136.82	3.29 × 10 <sup>-2</sup>							
	P2	136.84	4.31 × 10 <sup>-2</sup>							
	P1	136.85	9.86 × 10 <sup>-2</sup>							
	P1	136.87	8.15 × 10 <sup>-2</sup>							
	P1	136.90	5.98 × 10 <sup>-2</sup>							
	P1	136.91	5.58 × 10 <sup>-2</sup>							
	P1	136.92	6.02 × 10 <sup>-2</sup>							
	P1	136.95	0.11							
	P1	136.97	4.95 × 10 <sup>-2</sup>							
	P1	137.02	0.20		182 A/σ*(P-O)-Rydberg mixed					



Table 5. continued

peak/subpeak	calcd			EDGE	assignment	exptl
	core-hole site	$E$ -shifted <sup>b</sup> (eV)	$f \times 10^2$			$E$ (eV)
	P1	137.05	$5.44 \times 10^{-2}$		183 A/ $\sigma^*(\text{P-O})$ -Rydberg mixed	
	P2	137.10	$3.35 \times 10^{-2}$			

<sup>a</sup>Only the main transitions, computed at the TDDFT LB94 level, are reported. <sup>b</sup>Calculated excitation energies shifted by +2.2 eV to match the energy position of the first experimental peak.

is equal to 40 meV for TPPO (see Table 2), in good agreement with the measured value of 70 meV. A slightly higher value is found for PPT (50 meV compared with 30 meV derived from the experimental measurements).

**4.2. Assignment of the Experimental S and P L-Edge NEXAFS Spectra.** We start this discussion by assigning the measured absorption bands of the PPT S 2p NEXAFS spectrum to the dominant 1h-1p core excited configurations as obtained from the ZORA TDDFT LB94 calculations. The results are collected in Table 3 and displayed in Figure 4, where the experimental data are compared with the calculated profile. We report in the Supporting Information (Figure S2, upper panel) the NEXAFS spectrum of the DBT core of PPT taken from our recent work<sup>10</sup> in order to investigate the evolution of the spectrum in going from the DBT building block to PPT. The S 2p NEXAFS spectrum of PPT calculated by using the B3LYP xc potential is included in Figure S3 of the Supporting Information. Plots of selected virtual KS MOs are displayed in Figure S5. The complication inherent in the assignment of the S 2p core excitation spectrum is the presence of two series of transitions from the 2p SO core-hole components converging to the  $L_{\text{III}}$  and  $L_{\text{II}}$  thresholds, which overlap due to the small  $L_{\text{III}}-L_{\text{II}}$  SO splitting (which is around 1.2 eV). Further complexity is introduced by the redistribution of the intensity among the final excited states derived from the configuration mixing included in the TDDFT scheme. To facilitate the assignment, we applied a deconvolution of the computed spectra into the two series of transitions corresponding to the two SO components of the 2p hole states, as described in Section 3. In the discussion, we will refer to Table 3 where each band of the spectrum is associated with the character of the final MO involved in the dominant 1h-1p excitations contributing to each final excited state.

The total calculated S 2p NEXAFS spectrum of PPT is characterized by a low-lying weak spectral feature (peak A) which arises from several S  $2p_{3/2}$  transitions toward final MOs with a strong  $\pi^*(\text{C}=\text{C})$  valence antibonding character. These transitions have low intensity, consistently with the very small S *ns, nd* AO weights of the involved final MOs as well as with their high intensity observed in the C K-edge NEXAFS spectrum.<sup>9</sup> The corresponding transitions starting from the S  $2p_{1/2}$  level are centered at 165 eV (subpeak a', in blue) and contribute to the higher-energy portion of band B. The energy separation of a' and a'' subpeaks is 1.2 eV, the latter being close to both the calculated (1.31 eV) and the experimental (1.19 eV) values of the  $L_{\text{III}}/L_{\text{II}}$  SO splitting, thus confirming our assignment. The main component of peak B (subpeak b') arises from the  $L_{\text{III}}$  transitions toward final MOs with  $\sigma^*(\text{C-S})$  valence antibonding character with significant S diffuse s and d atomic components. Subpeak b'' contributes to the lower-energy side of the broad and strong band C which dominates the NEXAFS spectrum. Again, we observe a close match between the energy separation of the b' and b'' subpeaks (about 1.3 eV) and the experimental SO splitting. Subpeak c',

which is contributed by a manifold of transitions originating from the S  $2p_{3/2}$  levels, carries a large portion of the intensity of band C. The most intense transitions involve virtual MOs with mixed valence/Rydberg or pure Rydberg character with significant S *ns, nd* AO contributions. The valence character of these transitions involves both  $\sigma^*(\text{C-H})$  and  $\sigma^*(\text{P-O})$  MOs. The SO partner feature of subpeak c' (subpeak c'') contributes to the lower-energy side of peak D, while the higher-energy side (subpeak d') is associated with  $2p_{3/2}$  excitations toward virtual states of mixed valence/Rydberg character. Excitations of this nature also extend in the energy region of peak E, just below the  $L_{\text{III}}$  ionization thresholds. Since the SO partners of subpeaks d' and e' (subpeaks d'' and e'', respectively) are found just above the  $L_{\text{III}}$  ionization thresholds, they have been not included in Table 3. The good agreement between the TDDFT spectrum and the experimental one allows us to be confident of the proposed assignment.

Some strict similarities between the NEXAFS spectra of PPT and its DBT core emerge from a comparison of the upper and lower panels of Figure S2 (reported in the Supporting Information). In particular, the position of band B is relatively well-conserved in going from DBT to PPT, as well as the valence character,  $\sigma^*(\text{C-S})$ , of the virtual states. Moreover, both spectra are dominated by peak C, visible at around 166–167 eV, for which a similar assignment can be proposed, specifically to  $L_{\text{II}}$  transitions to virtual  $\sigma^*(\text{C-S})$  states and  $L_{\text{III}}$  transitions to virtual MOs of mixed valence-Rydberg character. In PPT, this peak is however also contributed by transitions to  $\sigma^*(\text{P-O})$  virtual states. At higher energy, a strict correspondence of the spectral nature and shape is hampered by the increasing number of overlapping transitions to diffuse MOs. To conclude this discussion, the fact that the most distinctive features of the NEXAFS spectra of DBT and PPT have a similar nature (energy position of the bands and their assignment) is a direct consequence of the localized nature of both initial and final states, which are rather insensitive to the addition of two phosphine oxide moieties in the more complex system of PPT.

Figure 5 reports the experimental P  $L_{\text{II,III}}$ -edge NEXAFS spectra of TPPO (upper panel) and PPT (lower panel), together with the TDDFT results obtained with the LB94 xc potential. Table 4 and Table 5 show the assignment of the experimental bands of TPPO and PPT, respectively, while the corresponding B3LYP results are reported in the Supporting Information (see Figure S4). Plots of selected KS virtual MOs are reported in Figure S6 and Figure S7 for TPPO and PPT, respectively. Also, in this case, the P 2p excitations give rise to two series of spectral features converging to the  $2p_{3/2}$  and  $2p_{1/2}$  ionization thresholds, which overlap due to the small SO energy splitting value (950 meV) of the  $L_{\text{II,III}}$  core-hole states. The assignment is more complicated in PPT than in TPPO because of the presence, in the former, of two nonequivalent P atoms as well as of the increased density of excited states.

Again, to facilitate the spectral attribution, for both molecules the computed spectra have been deconvoluted into the two series of transitions converging to the  $L_{\text{III}}$  and  $L_{\text{II}}$ -edges.

The calculated P  $L_{\text{II,III}}$ -edge NEXAFS spectrum of TPPO is characterized by a lower-energy double-peaked (A–B) structure: peaks A and B are contributed by the SO partner transitions from  $2p_{3/2}$  and  $2p_{1/2}$  core-holes toward low-lying virtual MOs of mainly  $\pi^*(\text{C}=\text{C})$  valence antibonding character. The two peaks are separated by a SO splitting of about 0.9 eV. Their low intensity reflects the small P  $nd$  AOs weights in the  $\pi^*$  valence MOs and indicates therefore that P 3d AOs are not involved in the formation of the P–O bond. Peak C is contributed by the  $2p_{3/2}$  transitions (subpeak b') toward final MOs with significant P diffuse  $s$  and  $d$  atomic components and some  $\sigma^*(\text{C}-\text{H})$  and  $\sigma^*(\text{P}-\text{O})$  valence antibonding characters. The  $L_{\text{II}}$  SO partner transitions (subpeak b'') contribute to peak D together with a manifold of  $L_{\text{III}}$  transitions (subpeak c') toward virtual MOs with a partial  $\sigma^*(\text{C}-\text{H})$  or  $\pi^*(\text{P}-\text{O})$  valence nature with considerable contributions of P diffuse  $s$ ,  $d$  AOs. Since the SO partner of subpeak c' (subpeak c'') is centered just above the  $L_{\text{III}}$  thresholds, its analysis is not included in Table 4.

A good correspondence between the energy position of the experimental and calculated peaks is observed, apart from the lacking, in the theoretical profile, of a weak structure centered between bands B and C. The intensity distribution is also qualitatively reproduced by theory, which is, however, not able to fully account for the finer structure of bands C and D, which we tentatively associate to vibronic effects not included in the computational protocol.

Similar considerations can be made in the case of PPT. With reference to the lower panel of Figure 5, the total calculated P  $L_{\text{II,III}}$ -edge NEXAFS spectrum of PPT is characterized, at lower excitation energies, by a weak feature (A) which corresponds to the (P1,P2)  $2p_{3/2}$  excitations toward final MOs of strong  $\pi^*(\text{C}=\text{C})$  valence antibonding character. The  $L_{\text{II}}$  SO partner transitions give rise to peak B (subpeak a''), which is also contributed by subpeak b', originating from the (P1,P2)  $2p_{3/2}$  excitations toward a final  $\pi^*$  MO. As in TPPO, the closeness between the energy separation of a' and a'' subpeaks (about 0.9 eV) and the experimental SO splitting (0.87 eV) confirms the assignments. Subpeak b'' is located in the lower-energy region of peak C; the latter mainly derives its intensity from a manifold of transitions converging to the  $L_{\text{III}}$  threshold (subpeak c'). The implicated virtual MOs present either mixed valence/Rydberg or pure Rydberg nature. Both  $\sigma^*(\text{P}-\text{O})$  and  $\sigma^*(\text{C}-\text{H})$  components contribute to the valence character of these transitions, in line with the spectral trend found for TPPO. Subpeak c'' is located under the envelope of band D, together with  $L_{\text{III}}$  transitions to final MOs of  $\sigma^*(\text{C}-\text{H})$  or  $\pi^*(\text{P}-\text{O})$  valence antibonding character together with sizable contributions from P  $ns$ ,  $nd$  AOs. In this energy region, weak transitions to diffuse orbitals are also visible. Subpeak d'' is centered at around 138 eV, just below the  $L_{\text{II}}$  threshold, and its analysis is not reported in Table 5. The main P 2p NEXAFS features are therefore conserved in going from TPPO to PPT, suggesting that the electronic behavior of the TPPO building block remains substantially unchanged when immersed in the more complex PPT.

An important point of this analysis concerns the relative energy position of the transitions toward  $\sigma^*(\text{P}-\text{O})$  and  $\pi^*(\text{P}-\text{O})$  states: the former occurs at lower energies compared to the latter (see Tables 4 and 5), while the  $\pi$  symmetry interaction in

the P–O bond is more stabilized compared to the  $\sigma$  one in the occupied MOs. These results therefore support the mechanism of negative hyperconjugation in the formation of the P–O bond<sup>41,42</sup> whereas the involvement of bent multiple bonds ( $\Omega$  bonds<sup>11</sup>) to the P–O bond is ruled out.

## 5. CONCLUSIONS

We present a joint experimental and theoretical investigation of the electronic structure of gas-phase TPPO and PPT as obtained through XPS and NEXAFS spectroscopies at the S and P  $L_{\text{II,III}}$ -edge region. The experimental results have been rationalized by means of the relativistic TDDFT approach which allows the inclusion of the coupling between 1h-1p excited configurations from the 2p degenerate core holes and gives a good account of the relativistic effects (mainly spin–orbit coupling) which are necessary to describe the transitions converging to the  $L_{\text{II}}$  and  $L_{\text{III}}$ -edges.

The calculation of the S 2p and P 2p XP spectra has allowed us to analyze the BEs in terms both of the SO splitting of the 2p core-holes and of the molecular-field splitting of the  $2p_{3/2}$  levels. A comparison of S 2p and P 2p XP spectra with those of the two building blocks of PPT, namely DBT and TPPO, reveals that both splittings are substantially conserved. The small increase of the S 2p experimental BEs in going from DBT to PPT is a consequence of the decreased shielding effect of the electronic charge density on sulfur due to the addition of two electron-withdrawing phosphine oxide moieties in PPT. The small decrease of the P  $L_{\text{III}}$  and  $L_{\text{II}}$  experimental BEs observed in going from TPPO to PPT can be instead rationalized by the replacement of a single phenyl ring of TPPO with one condensed ring of the DBT moiety in PPT.

The TDDFT results are accurate enough to provide a sound assignment of all absorption bands that characterize the below-threshold region of the S 2p and P 2p NEXAFS spectra. They display strong similarities to those of the PPT building blocks, DBT and TPPO, in line with the similar local environment of S and P atoms, being little affected by the increased molecular complexity of PPT. The present study complements the characterization of the electronic structure of TPPO and PPT obtained from the analysis of the C 1s and O 1s edge regions of these two molecules.<sup>9</sup> The intensity of the C 1s and O 1s NEXAFS spectra, which maps the C 2p and O 2p AO weights to the final MOs, derives from transitions to the low-lying delocalized  $\pi^*$  virtual valence orbitals, while the S 2p and P 2p NEXAFS intensities provide specific information on the higher-lying localized  $\sigma^*(\text{C}-\text{S})$  and  $\sigma^*(\text{P}-\text{O})$  virtual MOs. The results of this work indicate (i) that P 3d AOs are not involved in the formation of the P–O bond, and (ii) the energy ordering of P 2p transitions to  $\sigma^*(\text{P}-\text{O})$  and  $\pi^*(\text{P}-\text{O})$  virtual states is compatible with the traditional view of the P–O bond formation through a mechanism of negative hyperconjugation.<sup>11</sup> A similar study at the NEXAFS P K-edge would be useful to evaluate in detail the weight of the P 2p AO contributions to the P–O bond and will be the subject of future works.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcc.0c03973>.

Plots of selected MOs for TPPO and PPT; B3LYP XPS and NEXAFS spectra; VWN/TZP optimized Cartesian coordinates of DBT, TPPO, and PPT (PDF)

## AUTHOR INFORMATION

### Corresponding Authors

**D. Toffoli** – Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy;

orcid.org/0000-0002-8225-6119; Email: toffoli@units.it

**A. Guarnaccio** – ISM-CNR, Institute of Structure of Matter—, Tito Scalo (PZ) and Trieste, Italy; orcid.org/0000-0002-7927-5845; Email: ambra.guarnaccio@pz.ism.cnr.it

### Authors

**E. Bernes** – Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy

**G. Fronzoni** – Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy;

orcid.org/0000-0002-5722-2355

**M. Stener** – Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy;

orcid.org/0000-0003-3700-7903

**T. Zhang** – School of Information and Electronics, MIIT Key Laboratory for Low-Dimensional Quantum Structure and Devices, Beijing Institute of Technology (BIT), Beijing 100081, China; Department of Physics and Astronomy, Uppsala University, SE-751 20 Uppsala, Sweden

**C. Grazioli** – IOM-CNR, Laboratorio TASC, Sincrotrone Trieste, I-34149 Trieste, Basovizza, Italy

**F. O. L. Johansson** – Department of Physics and Astronomy, Uppsala University, SE-751 20 Uppsala, Sweden; orcid.org/0000-0002-6471-1093

**M. Coreno** – ISM-CNR, Institute of Structure of Matter—, Tito Scalo (PZ) and Trieste, Italy

**M. de Simone** – IOM-CNR, Laboratorio TASC, Sincrotrone Trieste, I-34149 Trieste, Basovizza, Italy

**C. Puglia** – Department of Physics and Astronomy, Uppsala University, SE-751 20 Uppsala, Sweden; orcid.org/0000-0001-6840-1570

Complete contact information is available at:  
<https://pubs.acs.org/10.1021/acs.jpcc.0c03973>

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work was supported by Università degli Studi di Trieste, Finanziamento di Ateneo per progetti di ricerca scientifica FRA 2015 and 2016. The authors thank the Carl Trygger Foundation for financial support and for making available the VG-Scienza SES-200 photoelectron analyzer at the Gas Phase beamline, Elettra, Italy. The authors acknowledge the EU CERIC–ERIC Consortium for the access to experimental facilities and financial support. T.Z. is grateful for the financial support from the NSFC (Grant No. 61901038 and 61971035) and the support from the Beijing Institute of Technology Research Fund Program for Young Scholars. F.J. acknowledges the support from the Swedish Research Council (Grant 2014-6463) and Marie Skłodowska-Curie Actions (Cofund, Project INCA 600398).

## REFERENCES

- (1) Adachi, C.; Baldo, M. A.; Thompson, M. E.; Forrest, S.R. J. Nearly 100% internal phosphorescence efficiency in an organic light-emitting device. *J. Appl. Phys.* **2001**, *90*, 5048–5051.
- (2) Tang, C. W.; Van Slyke, S. A. Organic Electroluminescent Diodes. *Appl. Phys. Lett.* **1987**, *51*, 913–915.
- (3) Friend, R. H.; Gymer, R. W.; Holmes, A. B.; Burroughes, J. H.; Marks, R. N.; Taliani, C.; Bradley, D. D. C.; Dos Santos, D. A.; Brédas, J.-L.; Lögdlund, M.; et al. Electroluminescence in conjugated polymers. *Nature* **1999**, *397*, 121–128.
- (4) Kim, D.; Salman, S.; Coropceanu, V.; Salomon, E.; Padmaperuma, A. B.; Sapochak, L. S.; Kahn, A.; Brédas, J.-L. Phosphine Oxide Derivatives as Hosts for Blue Phosphors: A Joint Theoretical and Experimental Study of Their Electronic Structure. *Chem. Mater.* **2010**, *22*, 247–254.
- (5) Tsai, Y.-S.; Hong, L.-A.; Juang, F.-S.; Chen, C.-Y. Blue and white phosphorescent organic light emitting diode performance improvement by confining electrons and holes inside double emitting layers. *J. Lumin.* **2014**, *153*, 312–316.
- (6) Chapran, M.; Angioni, E.; Findlay, N. J.; Breig, B.; Cherpak, V.; Stakhira, P.; Tuttle, T.; Volyniuk, D.; Grazulevicius, J. V.; Nastishin, Y. A.; et al. An Ambipolar BODIPIY Derivative for a White Exciplex OLED and Cholesteric Liquid Crystal Laser toward Multifunctional Devices. *ACS Appl. Mater. Interfaces* **2017**, *9*, 4750–4757.
- (7) Sapochak, L. S.; Padmaperuma, A. B.; Vecchi, P. A.; Cai, X.; Burrows, P. E. Designing Organic Phosphine Oxide Host Materials Using Heteroaromatic Building Blocks: Inductive Effects on Electroluminescence. *Proc. SPIE* **2007**, 665506.
- (8) Stöhr, J. *NEXAFS Spectroscopy*; Springer Verlag: Berlin, Heidelberg, Germany, 1992.
- (9) Guarnaccio, A.; Zhang, T.; Grazioli, C.; Johansson, F.; Coreno, M.; de Simone, M.; Fronzoni, G.; Toffoli, D.; Bernes, E.; Puglia, C. PPT Isolated Molecule and Its Building Block Moieties Studied by C 1s and O 1s Gas Phase X-ray Photoelectron and Photoabsorption Spectroscopies. *J. Phys. Chem. C* **2020**, *124*, 9774–9786.
- (10) Toffoli, D.; Guarnaccio, A.; Grazioli, C.; Zhang, T.; Johansson, F.; de Simone, M.; Coreno, M.; Santagata, A.; D’Auria, M.; Puglia, C.; et al. Electronic Structure Characterization of a Thiophene Benzo-Annulated Series of Common Building Blocks for Donor and Acceptor Compounds Studied by Gas Phase Photoelectron and Photoabsorption Synchrotron Spectroscopies. *J. Phys. Chem. A* **2018**, *122*, 8745–8761.
- (11) Gilheany, D. G. No d Orbitals but Walsh Diagrams and Maybe Banana Bonds: Chemical Bonding in Phosphines, Phosphine Oxides, and Phosphonium Ylides. *Chem. Rev.* **1994**, *94*, 1339–1374.
- (12) Blyth, R.; Delaunay, R.; Zitnik, M.; Krempasky, J.; Krempaska, R.; Slezak, J.; Prince, K. C.; Richter, R.; Vondracek, M.; Camilloni, R.; et al. The high resolution Gas Phase Photoemission beamline, Elettra. *J. Electron Spectrosc. Relat. Phenom.* **1999**, *101–103*, 959–964.
- (13) Mårtensson, N.; Baltzer, P.; Brühwiler, P. A.; Forsell, J.-O.; Nilsson, A.; Stenborg, A.; Wannberg, B. A Very High Resolution Electron Spectrometer. *J. Electron Spectrosc. Relat. Phenom.* **1994**, *70*, 117–128.
- (14) Pettersson, L.; Nordgren, J.; Selander, L.; Nordling, C.; Siegbahn, K.; Ågren, H. Core-Electron Binding Energies in The Soft X-Ray Range Obtained in X-Ray Emission. *J. Electron Spectrosc. Relat. Phenom.* **1982**, *27*, 29–37.
- (15) True, J. E.; Thomas, T. D.; Winter, R. W.; Gard, G. L. Electronegativities from Core-Ionization Energies: Electronegativities of SF<sub>5</sub> and CF<sub>3</sub>. *Inorg. Chem.* **2003**, *42*, 4437–4441.
- (16) Eustatii, I. G.; Francis, J. T.; Tyliczszak, T.; Turci, C. C.; Kilcoyne, A. L. D.; Hitchcock, A. P. Generalized Oscillator Strengths for Inner-Shell Excitation of SF<sub>6</sub> Recorded with a High-Performance Electron Energy Loss Spectrometer. *Chem. Phys.* **2000**, *257*, 235–252.
- (17) Haensel, R.; Keitel, G.; Schreiber, P.; Kunz, C. Optical Absorption of Solid Krypton and Xenon in the Far Ultraviolet. *Phys. Rev.* **1969**, *188*, 1375–1380.
- (18) Giertz, A.; Bäessler, M.; Björneholm, O.; Wang, H.; Feifel, R.; Miron, C.; Karlsson, L.; Svensson, S.; Børve, K. J.; Sæthre, L. J. High

Resolution C1s and S2p Photoelectron Spectra of Thiophene. *J. Chem. Phys.* **2002**, *117*, 7587–7592.

(19) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.

(20) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: a Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.

(21) Fonseca Guerra, C.; Snijders, J. G.; teVelde, G.; Baerends, E. J. Towards an Order-N DFT Method. *Theor. Chem. Acc.* **1998**, *99*, 391–403.

(22) 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.

(23) Bandoli, G.; Bortolozzo, G.; Clemente, D. A.; Croatto, U.; Panattoni, C. Crystal and Molecular Structure of Triphenylphosphine Oxide. *J. Chem. Soc. A* **1970**, 2778–2780.

(24) Al-Farhan, K. A. Crystal structure of triphenylphosphine oxide. *J. Crystallogr. Spectrosc. Res.* **1992**, *22*, 687–689.

(25) Diop, C. A. K.; Touré, A.; Diop, L.; Welter, R. 2-Benzoylbenzoato- kO triphenyl (triphenylphosphineoxide-kO) tin(IV). *Acta Crystallogr., Sect. E: Struct. Rep. Online* **2006**, E62, m3338–m3340.

(26) De Silva, N.; Zahariev, F.; Hay, B. P.; Gordon, M. S.; Windus, T. L. Conformations of Organophosphine Oxides. *J. Phys. Chem. A* **2015**, *119*, 8765–8773.

(27) Wang, F.; Ziegler, T.; van Lenthe, E.; van Gisbergen, S.; Baerends, E. J. The Calculation of Excitation Energies Based on The Relativistic Two-Component Zeroth-Order Regular Approximation and Time-Dependent Density-Functional with Full Use of Symmetry. *J. Chem. Phys.* **2005**, *122*, 204103.

(28) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J.A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.

(29) Stener, M.; Fronzoni, G.; de Simone, M. Time Dependent Density Functional Theory of Core Electrons Excitations. *Chem. Phys. Lett.* **2003**, *373*, 115–123.

(30) Fronzoni, G.; Stener, M.; Decleva, P.; Wang, F.; Ziegler, T.; van Lenthe, E.; Baerends, E. J. Spin-Orbit Relativistic Time Dependent Density Functional Theory Calculations for The Description of Core Electron Excitations: TiCl4 Case Study. *Chem. Phys. Lett.* **2005**, *416*, 56–63.

(31) Casida, M. E. Time-Dependent Density Functional Response Theory for Molecules. *Recent Advances in Density Functional Methods* **1995**, *1*, 155–192.

(32) 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.

(33) Cederbaum, L. S.; Domcke, W.; Schirmer, J. Many-Body Theory of Core Holes. *Phys. Rev. A: At., Mol., Opt. Phys.* **1980**, *22*, 206–222.

(34) Barth, A.; Cederbaum, L. S. Many-Body Theory of Core-Valence Excitations. *Phys. Rev. A: At., Mol., Opt. Phys.* **1981**, *23*, 1038–1061.

(35) Decleva, P.; Fronzoni, G.; Lisini, A.; Stener, M. Molecular Orbital Description of Core Excitation Spectra in Transition Metal Compounds. *Chem. Phys.* **1994**, *186*, 1–16.

(36) van Lenthe, E.; Baerends, E. J. Optimized Slater-Type Basis Sets for The Elements 1–118. *J. Comput. Chem.* **2003**, *24*, 1142–1156.

(37) van Leeuwen, R.; Baerends, E. J. Exchange-Correlation Potential with Correct Asymptotic Behaviour. *Phys. Rev. A: At., Mol., Opt. Phys.* **1994**, *49*, 2421–2431.

(38) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(39) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.

(40) 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.* **1994**, *98*, 11623–11627.

(41) Alabugin, I. V.; Gilmore, K. M.; Peterson, P. W. Hyperconjugation. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 109–141.

(42) Reed, A. E.; Schleyer, P. v. R. Chemical bonding in hypervalent molecules. The dominance of ionic bonding and negative hyperconjugation over d-orbital participation. *J. Am. Chem. Soc.* **1990**, *112*, 1434–1445.

# Supporting Information for: S 2p and P 2p Core Level Spectroscopy of PPT Ambipolar Material and Its Building Block Moieties

E. Bernes,<sup>a</sup> G. Fronzoni,<sup>a</sup>, M. Stener,<sup>a</sup> A. Guarnaccio,<sup>b</sup> T. Zhang,<sup>c,d</sup> C. Grazioli,<sup>e</sup>

F. O. L. Johansson,<sup>d</sup> M. Coreno,<sup>b</sup> M. de Simone,<sup>e</sup> C. Puglia,<sup>\*,d</sup> D. Toffoli,<sup>\*,a</sup>

<sup>a</sup>Department of Chemical and Pharmaceutical Sciences, University of Trieste, 34127 Trieste, Italy

<sup>b</sup>ISM-CNR, Institute of Structure of Matter -Tito Scalo (PZ) and Trieste, Italy

<sup>c</sup>School of Information and Electronics, MIIT Key Laboratory for Low-Dimensional Quantum Structure and Devices, Beijing Institute of Technology (BIT), Beijing 100081, China

<sup>d</sup>Department of Physics and Astronomy, Uppsala University, Box 516, SE-751 20 Uppsala, Sweden

<sup>e</sup>IOM-CNR, Laboratorio TASC, Sincrotrone Trieste, I-34149 Trieste, Basovizza, Italy

## Table of Content

<b>Table S1:</b> .....	2
<b>Table S2:</b> .....	2
<b>Fig. S1:</b> .....	3
<b>Fig. S2:</b> .....	4
<b>Fig. S3:</b> .....	5
<b>Fig. S4:</b> .....	6
<b>Fig. S5:</b> .....	7
<b>Fig. S6:</b> .....	7
<b>Fig. S7:</b> .....	8
<b>Optimized coordinates</b> .....	9-11

**Table S1.** Comparison between experimental and theoretical S2p IPs for PPT and DBT. Theoretical IPs are obtained by using the B3LYP and LB94 xc potentials for PPT and DBT, respectively. Data for the latter are taken from ref.<sup>1</sup>. See main text for the details about the basis sets used in the TDDFT calculations. All values are expressed in eV.

Edge	PPT		DBT	
	Theory <sup>a</sup>	Single fit	Theory <sup>b</sup>	Single fit
<b>L<sub>II</sub></b>	170.69	170.59	170.57	170.47
<b>L<sub>III</sub></b>	169.46	169.46	169.19	169.16
<b>L<sub>III</sub></b>	169.35	169.34	169.33	169.32

<sup>a</sup>Calculated IPs shifted by +8.23 eV.

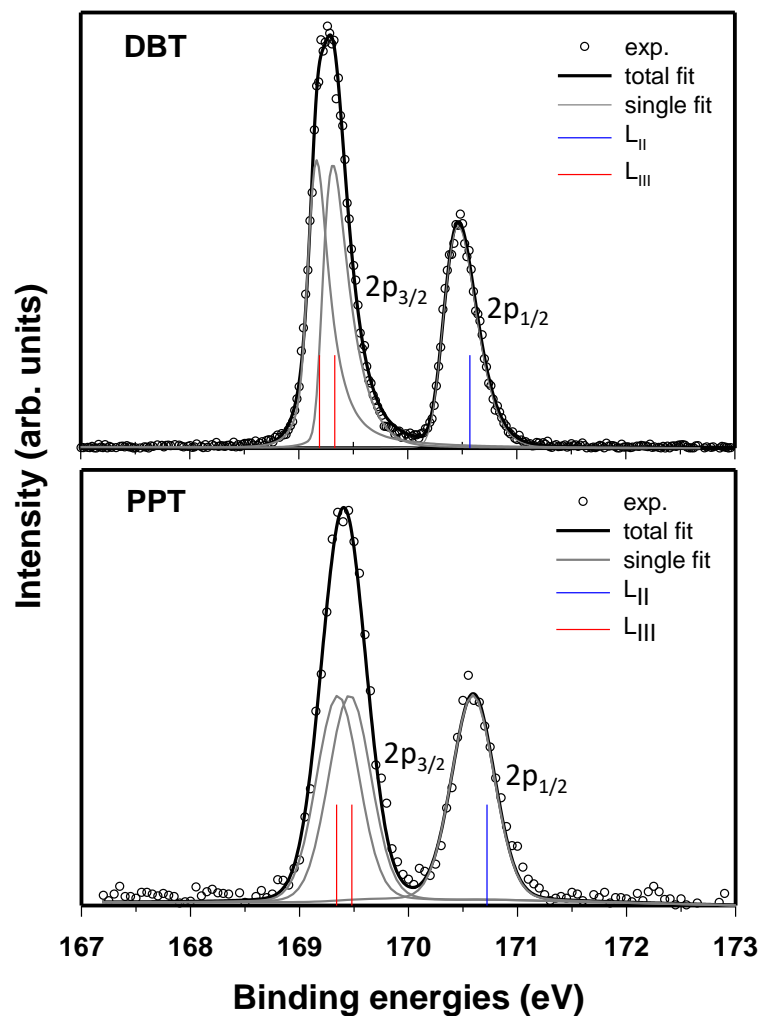
<sup>b</sup>Calculated IPs shifted by -0.90 eV.

**Table S2.** Comparison between experimental and theoretical P2p IPs of TPPO and PPT. Theoretical IPs are obtained by using the B3LYP xc potential. See main text for the details about the basis sets used in the TDDFT calculations. All values are expressed in eV.

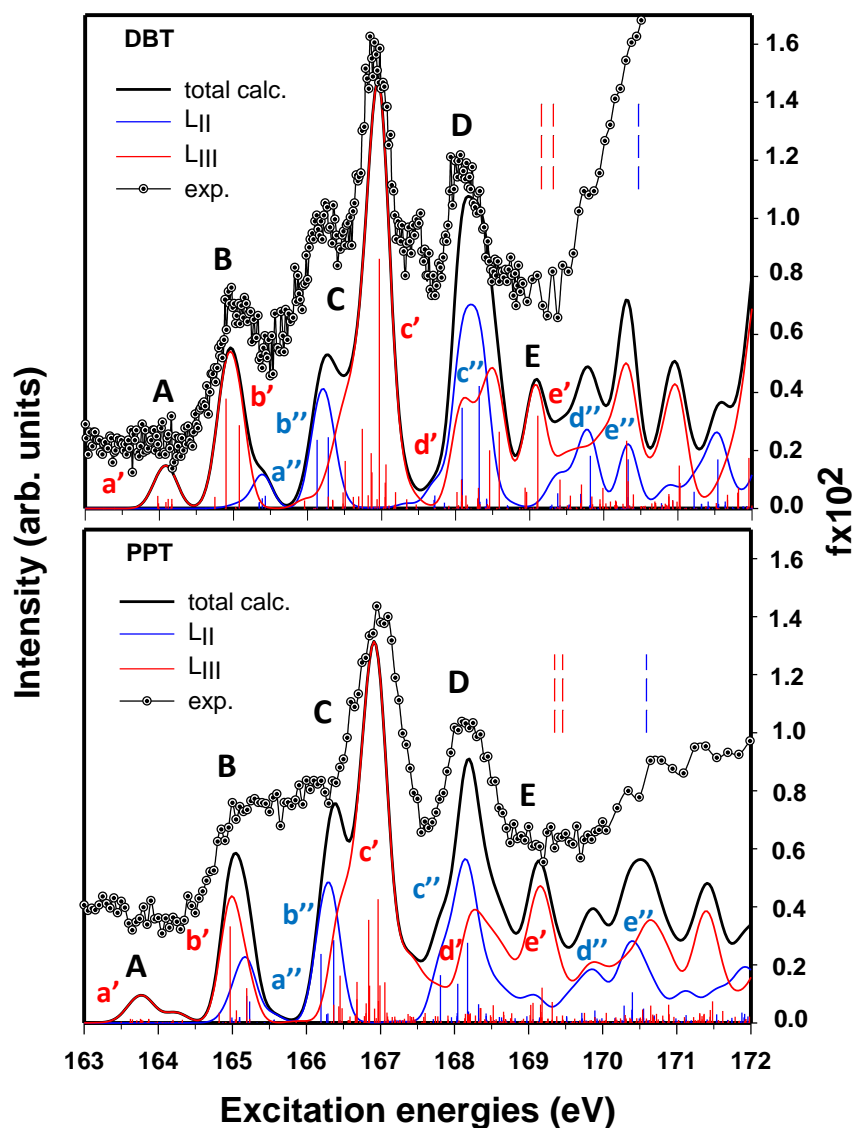
Edge	TPPO		PPT		
	Theory <sup>a</sup>	Single fit	Theory <sup>b</sup>		Single fit
			P1	P2	
<b>L<sub>II</sub></b>	138.21	138.15	138.22	138.14	138.01
<b>L<sub>III</sub></b>	137.28	137.28	137.33	137.32	137.22
<b>L<sub>III</sub></b>	137.27	137.27	137.22	137.20	137.10

<sup>a</sup>Calculated IPs shifted by +8.17 eV.

<sup>b</sup>Calculated IPs shifted by +7.91 eV.

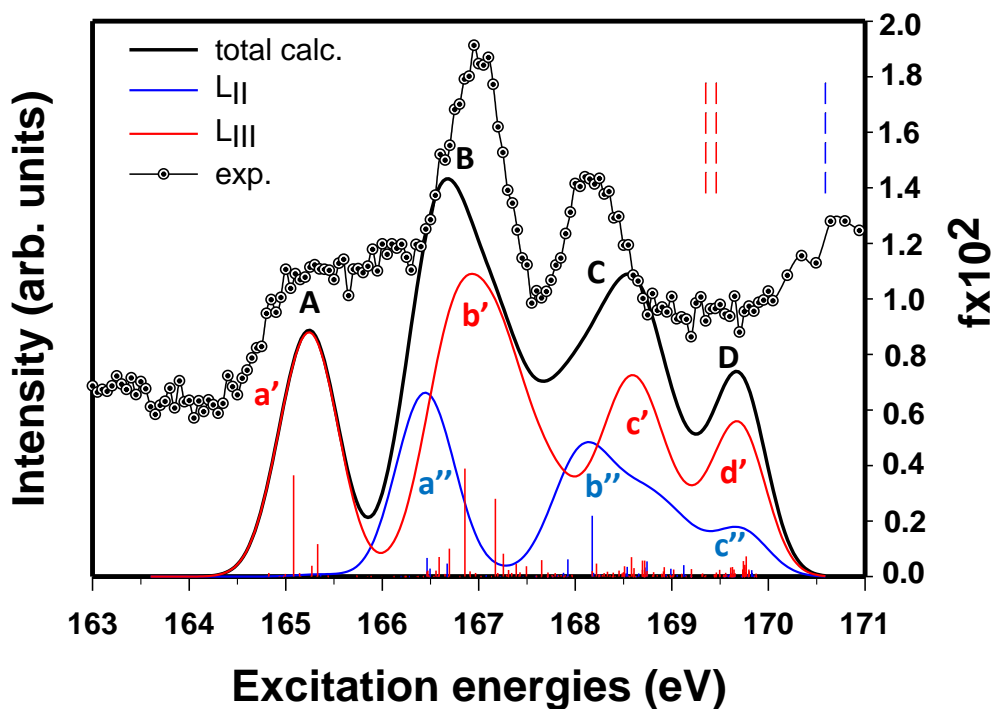


**Figure S1.** S<sub>2p</sub> XP spectra of DBT (upper panel) and PPT (lower panel) calculated by employing the LB94 xc potential: experimental data (circles) are shown together with the results of the total fit (black line) obtained by the procedure described in the text. The vertical colored bars are the theoretical IPs and have been shifted by  $-0.90$  eV and  $-1.23$  eV for DBT and PPT, respectively. The light gray curves are the L<sub>II</sub> (centered around the blue vertical bar) and L<sub>III</sub> (centered around the red vertical bars) S<sub>2p</sub> components obtained by the fitting procedure using a single and two asymmetric Voigt functions, respectively.

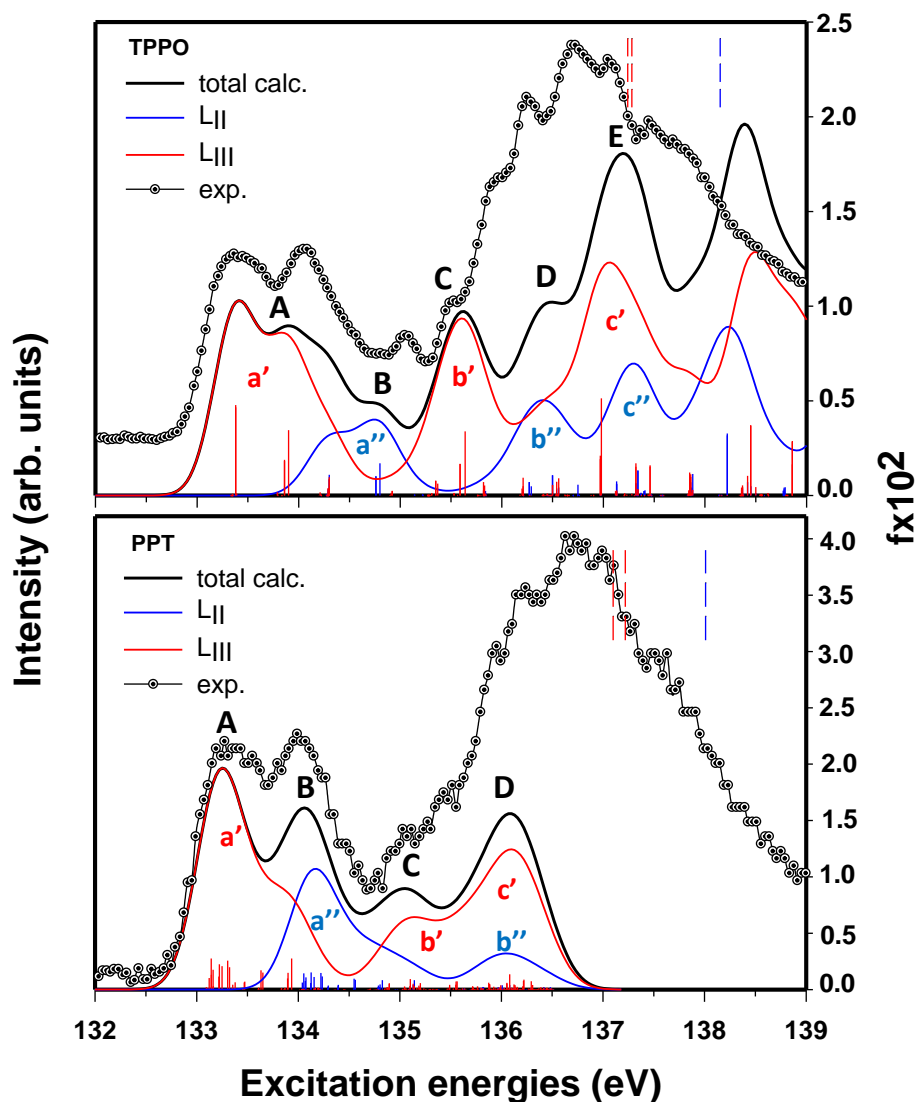


**Figure S2.** S  $L_{II,III}$ -edge NEXAFS spectra of DBT (upper panel) and PPT (lower panel) calculated by employing the LB94 xc potential: experimental data (circles), calculated TDDFT results (black solid line). Also shown is the deconvolution of the calculated S2p spectrum into the two manifolds of excited states converging to the  $L_{III}$  (red solid line and vertical red bars) and  $L_{II}$  (blue solid line and vertical blue bars) edges. The energy scale of the calculated data has been shifted by +0.4 eV and +0.3 eV, respectively for DBT and PPT in order to match the first experimental peak. The experimental S2p ionization thresholds are also shown (blue and red vertical bars). The left vertical axis refers to the experimental intensities (plotted in arbitrary units), while the right vertical axis refers to computed oscillator strengths (multiplied by a factor of 100).

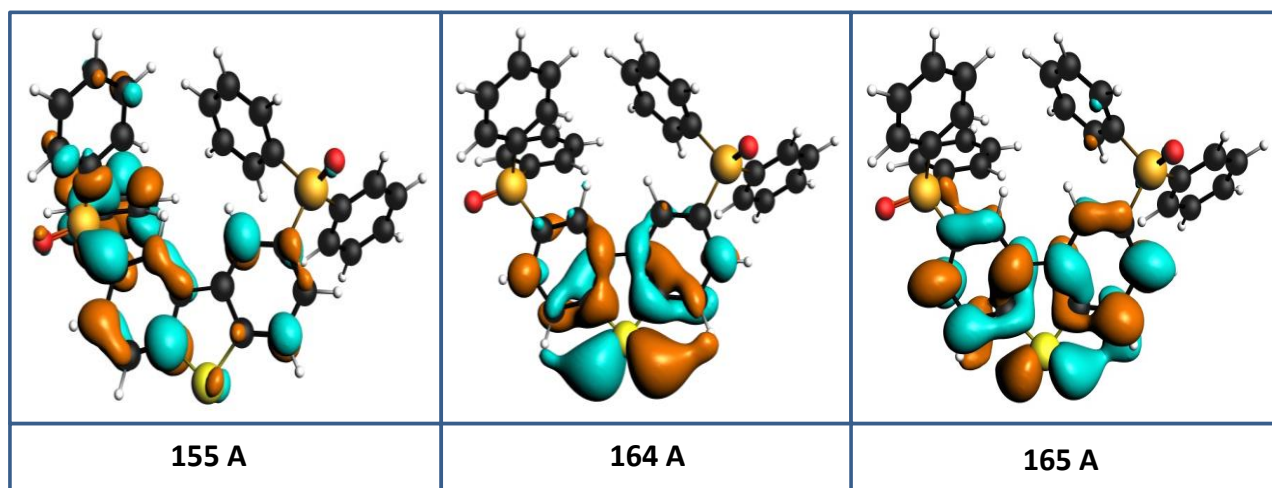




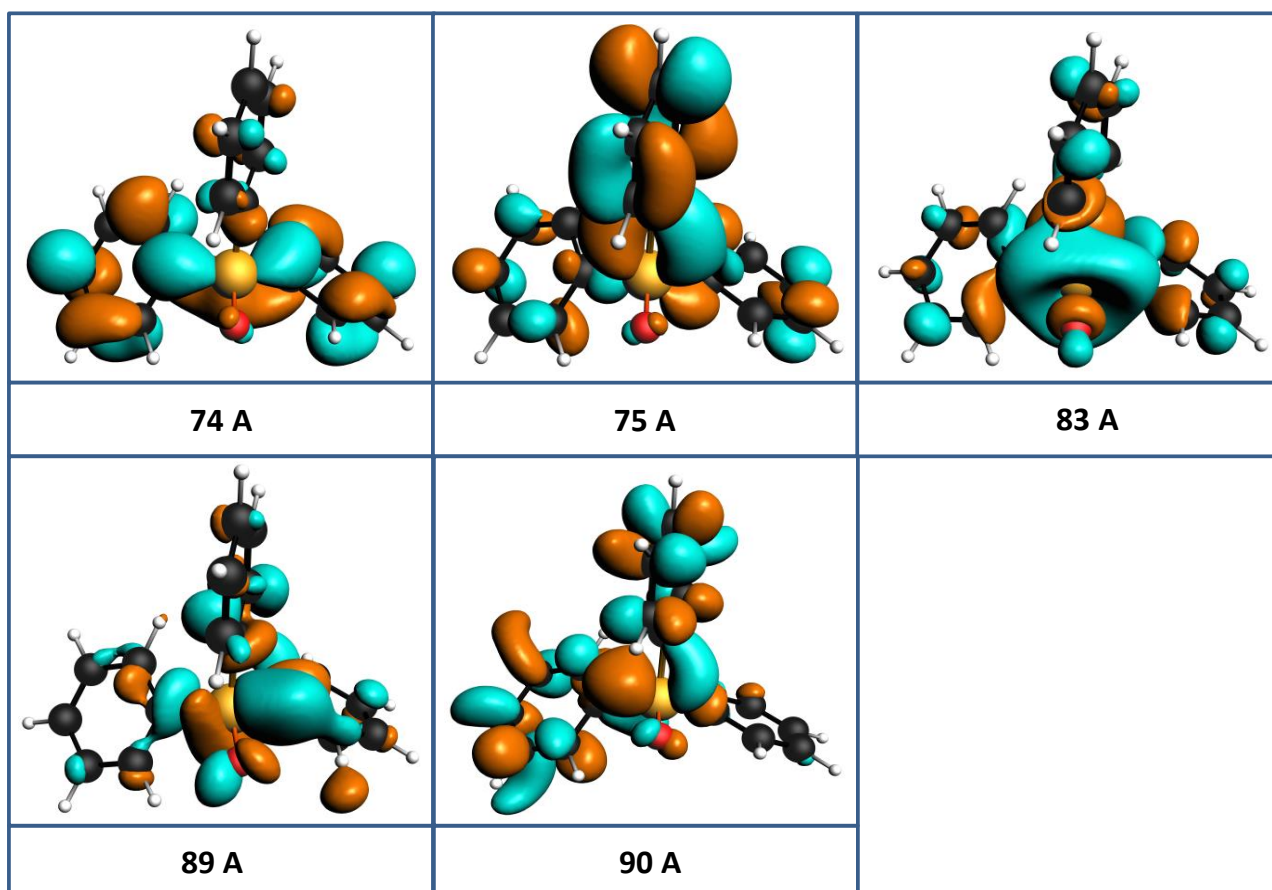
**Figure S3.** S  $L_{II,III}$ -edge NEXAFS spectrum of PPT calculated by employing the hybrid B3LYP xc potential: experimental data (circles), calculated TDDFT results (black solid line). Also shown is the deconvolution of the calculated S2p spectrum into the two manifolds of excited states converging to the  $L_{III}$  (red solid line and vertical red bars) and  $L_{II}$  (blue solid line and vertical blue bars) edges. The energy scale of the calculated data has been shifted by +0.3 eV, in order to match the first experimental peak. The experimental S2p ionization thresholds are also shown (blue and red vertical bars). The left vertical axis refers to the experimental intensities (plotted in arbitrary units), while the right vertical axis refers to computed oscillator strengths (multiplied by a factor of 100).



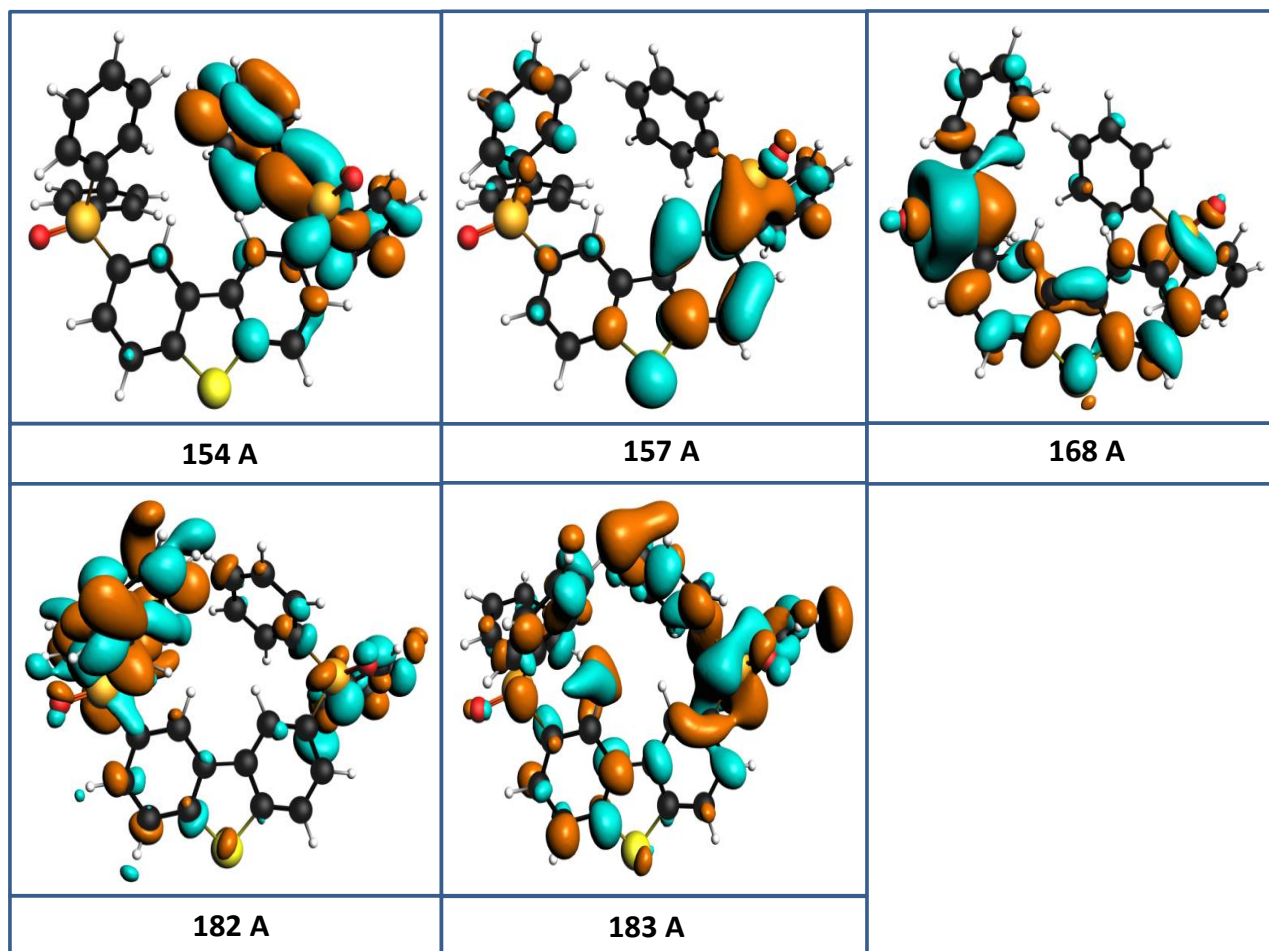
**Figure S4.** P  $L_{II,III}$ -edge NEXAFS spectra of TPPO (upper panel) and PPT (lower panel) calculated by employing the hybrid B3LYP xc potential: experimental data (circles), calculated TDDFT results (black solid line). Also shown is the deconvolution of the calculated P2p spectrum into the two manifolds of excited states converging to the  $L_{III}$  (red solid line and vertical red bars) and  $L_{II}$  (blue solid line and vertical blue bars) edges. The energy scale of the calculated data has been shifted by +6.7 eV and +6.6 eV, respectively for TPPO and PPT in order to match the first experimental peak. The experimental P2p ionization thresholds are also shown (blue and red vertical bars). The left vertical axis refers to the experimental intensities (plotted in arbitrary units), while the right vertical axis refers to computed oscillator strengths (multiplied by a factor of 100).



**Figure S5.** Plots of selected KS virtual molecular orbitals for PPT. For all atoms except S, an all-electron DZP basis set is used. A ET-QZ4P-2DIFFUSE basis is used for S. See main text for details about the basis set used in the TDDFT calculations.



**Figure S6.** Plots of selected KS virtual molecular orbitals for TPPO. For all atoms except P, an all-electron DZP basis set is used. A ET-QZ4P-2DIFFUSE basis is used for P. See main text for details about the basis set used in the TDDFT calculations.



**Figure S7.** Plots of selected KS virtual molecular orbitals for PPT. For all atoms except P, an all-electron DZP basis set is used. A ET-QZ4P-2DIFFUSE basis is used for P. See main text for details about the basis set used in the TDDFT calculations.

## Optimized Cartesian coordinates (in Å) of DBT, TPPO, and PPT.

### DBT

S	0.165399	2.327192	0.000000
C	-0.727585	-0.129278	0.000000
C	0.703081	-0.230696	0.000000
C	-1.160065	1.206449	0.000000
C	1.319853	1.030856	0.000000
C	-1.679909	-1.143507	0.000000
C	1.500988	-1.370201	0.000000
C	-2.507545	1.533733	0.000000
C	2.700442	1.162621	0.000000
C	-3.022341	-0.823855	0.000000
C	2.875379	-1.245191	0.000000
C	-3.433981	0.508731	0.000000
C	3.471033	0.015890	0.000000
H	-1.358014	-2.189725	0.000000
H	1.032966	-2.359716	0.000000
H	-2.828436	2.578875	0.000000
H	3.167404	2.151159	0.000000
H	-3.771151	-1.620621	0.000000
H	3.503630	-2.140051	0.000000
H	-4.500771	0.748042	0.000000
H	4.561032	0.101408	0.000000

### TPPO

P	-2.112017	0.995541	-1.057376
O	-2.118779	1.069084	-2.546294
C	-3.531391	1.810449	-0.321878
C	-0.657460	1.759385	-0.335448
C	-2.137326	-0.693132	-0.448509
C	-4.195507	2.728631	-1.126724
C	-5.301807	3.401281	-0.635924
C	-5.743352	3.159399	0.656643
C	-5.086596	2.235672	1.457421
C	-3.983609	1.555990	0.968390
C	0.435542	1.913688	-1.180239
C	-0.586349	2.188345	0.985459
C	-2.584066	-1.662567	-1.338917
C	-2.639311	-2.988499	-0.943441
C	-1.732482	-1.054216	0.832183
C	-2.251309	-3.346186	0.339564
C	-1.795832	-2.380429	1.225845
C	0.585062	2.751070	1.464214
C	1.602628	2.482600	-0.699186
C	1.679482	2.894704	0.623116
H	-5.830704	4.117214	-1.271641
H	-6.621757	3.686632	1.041159
H	-5.449590	2.029669	2.468347
H	-3.833948	2.880307	-2.149494
H	-3.492709	0.790434	1.580710
H	2.600585	3.349121	1.000669
H	0.639939	3.097205	2.500215
H	2.460482	2.612416	-1.365264

H	0.335628	1.595877	-2.223652
H	-1.327088	-0.296648	1.513125
H	-1.468693	-2.667126	2.229324
H	-2.289462	-4.395078	0.649136
H	-2.981941	-3.753367	-1.646274
H	-2.863524	-1.350418	-2.350907
H	-1.467815	2.116631	1.633414

**PPT**

C	-5.880226	-3.390747	3.920101
C	-0.379390	1.770507	3.305118
C	2.764076	4.108057	0.337233
C	3.894865	-1.135051	5.343652
C	-6.466789	-2.236315	3.422107
C	-4.585775	-3.731814	3.553459
C	-0.562401	0.434917	3.643024
C	-1.233405	2.380785	2.399361
C	2.068049	3.244082	1.170341
C	3.922121	3.680983	-0.295855
C	2.765457	-1.164650	4.537790
C	5.130648	-0.827826	4.794434
C	-5.757563	-1.415261	2.561759
C	-3.874297	-2.916213	2.689349
C	-1.582892	-0.293393	3.056810
C	-2.250192	1.650290	1.802155
C	2.526656	1.954403	1.371963
C	4.380549	2.388004	-0.104829
C	2.868245	-0.883690	3.186110
C	5.237436	-0.539238	3.443507
C	-2.990240	-2.746635	-0.597241
C	3.887604	-2.273208	-0.745764
C	-2.127038	-3.551671	-1.317416
C	3.112322	-3.200915	-1.417395
C	-1.170094	-1.329040	0.080082
C	1.922611	-1.095366	0.006404
C	-4.457672	-1.750963	2.201047
C	-2.415310	0.309416	2.118781
C	3.679623	1.520881	0.725260
C	4.105592	-0.560246	2.636933
C	-2.516186	-1.638590	0.120775
C	3.297091	-1.227726	-0.027262
C	-0.289410	-2.139580	-0.608779
C	1.132511	-2.039225	-0.628298
C	-0.766490	-3.257490	-1.302664
C	1.729119	-3.089376	-1.339906
O	-4.533993	0.299945	0.392647
O	5.791589	-0.218219	0.493656
P	-3.602590	-0.608582	1.116630
P	4.362925	-0.129251	0.907061
S	0.535151	-4.164732	-2.013752
H	-6.443030	-4.040109	4.597728
H	-7.491340	-1.976657	3.703565
H	-4.131229	-4.650730	3.934358
H	-6.196161	-0.505630	2.137915
H	-2.864902	-3.198724	2.367862
H	0.437855	2.339249	3.761234

H	0.107348	-0.038905	4.368042
H	-1.102883	3.437006	2.145455
H	-1.724282	-1.349768	3.312302
H	-2.927102	2.093658	1.063834
H	-4.057464	-2.988349	-0.572218
H	-2.508096	-4.418368	-1.862642
H	-0.780588	-0.464213	0.626232
H	2.404284	5.130524	0.186872
H	1.157492	3.577907	1.678242
H	4.477555	4.365732	-0.943047
H	1.991409	1.293235	2.065169
H	5.300792	2.027095	-0.576609
H	3.811846	-1.367909	6.409643
H	1.794336	-1.433250	4.966277
H	6.023825	-0.818481	5.425732
H	1.977360	-0.949208	2.550442
H	6.201769	-0.306419	2.978849
H	4.980264	-2.332447	-0.763615
H	3.581196	-4.011436	-1.980426
H	1.449630	-0.249858	0.519658

## References

---

<sup>1</sup> Toffoli, D.; Guarnaccio, A.; Grazioli, C.; Zhang, T.; Johansson, F.; de Simone, M.; Coreno, M.; Santagata, A.; D'Auria, M.; Puglia, C. et al. Electronic Structure Characterization of a Thiophene Benzo-Annulated Series of Common Building Blocks for Donor and Acceptor Compounds Studied by Gas Phase Photoelectron and Photoabsorption Synchrotron Spectroscopies. *J. Phys. Chem. A* **2018**, *122*, 8745–8761

# A Theoretical and Experimental Electronic Structure Characterization of Titanyl Phthalocyanine by means of Angle-Resolved Near-Edge X-ray Absorption Fine Structure Spectroscopy at C,N, and O K-Edges

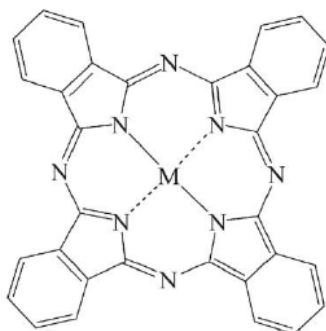
## Introduction

During the past decades, great efforts have been made in the field of organic semiconductors (OSCs) due to their peculiar electronic and optical properties which allow them to be implemented in molecular electronics devices, such as organic photovoltaic cells (OPVCs) and organic light emitting diodes (OLEDs)<sup>1,2</sup>. OSCs typically form well-ordered thin films on different kind of substrates; however, the preparation of ordered OSC films for optoelectronic applications represents a challenge due to the fact that optoelectronic properties of OSC solids are rather anisotropic and affected by the molecular and crystalline orientation<sup>3,4</sup> as well as on the polymorphs<sup>5,6</sup>. Further difficulties are represented by the small sublimation enthalpies and low barriers of diffusion processes<sup>7,8</sup> due to the weak molecule-substrate interaction typical of OSCs.

To exploit the full potential of OSCs for electronic and optoelectronic devices, it is fundamental to get a detailed knowledge on both structural and electronic properties of metal-organic interfaces and adsorbate systems, especially electronic interactions and growth mechanisms. One of the most promising classes of OSCs is represented by phthalocyanines (Pcs), a group of macrocyclic molecular complexes which possess an extended  $\pi$ -system due to the presence of four pyrrolic rings. The N atoms at the interior form a central cavity with two acidic H atoms which can be easily replaced by a metal cation in a tetradentate fashion, thus forming a metal Pc (MPc). More than 70 different ions are known to be incorporated in the central pocket; the nature of the metal center strongly affects the functional properties of the resulting molecule, such as the regulation of vital processes (*e.g.*, electron transfer, reversible axial ligation of adducts, light-harvesting and catalytic transformations). Moreover, tailoring of substituent groups on the macrocycles allows to optimize the performance of these systems, determining their arrangement in different environments as well as the assembly of molecular nanoarchitectures<sup>9,10</sup>. MPcs are particularly well exploited as blue-green synthetic pigments for printing inks, paints, and color filters<sup>11</sup>. Furthermore, these complexes are also employed as photosensitizers in dye sensitized solar cells, photoconductors, organic solar cells, electro-chromic materials, photovoltaic cells, and for photo-dynamic cancer therapy<sup>12,13</sup> thanks to their high photostability and strong light absorption in the visible range. Moreover, MPcs play fundamental roles in nature, such as in oxygen transport, electron transfer, oxidation catalysis and photosynthesis.

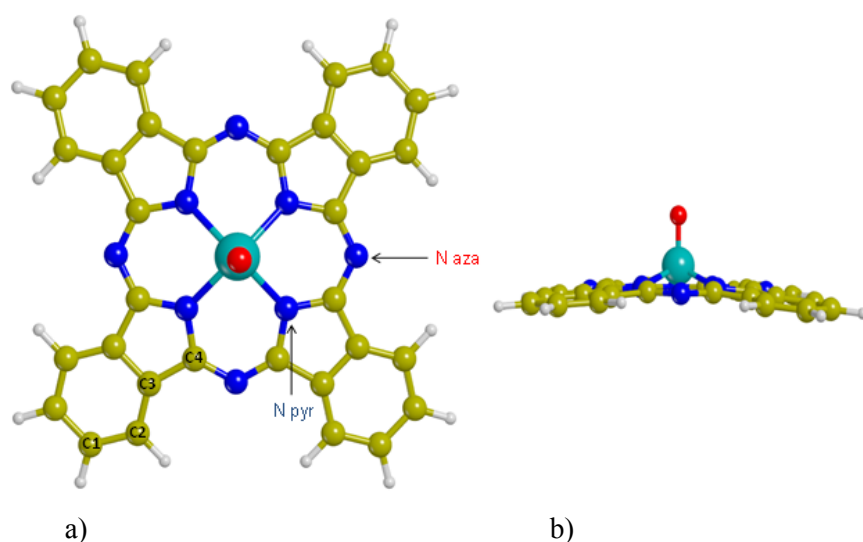


Generally, MPcs are planar; a generic chemical structure is reported in Figure 1. However, Pc complexes can also assume a non-planar structure, when the size of the metal ion exceeds the space available in the center of the molecule, as in the case of Sn, Pb and the lanthanoids, or when an axial ligand binds to the metal center, as in titanyl phthalocyanine (TiOPc), the MPc investigated in this thesis work.



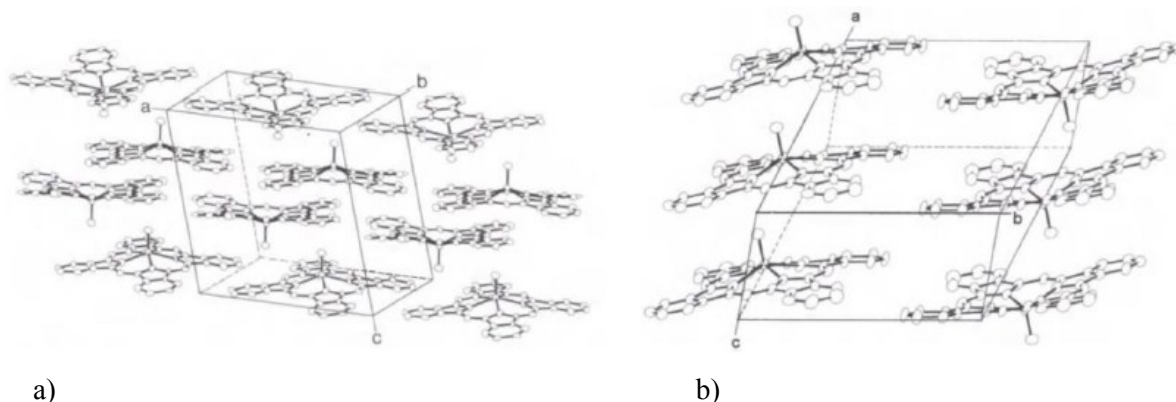
**Figure 1** – *Chemical structure of a generic MPc.*

The electronic structure of these complexes is affected by the nature of the metal present in the central cavity, which determines their reactivity as well as their chemical and physical properties. In particular, these complexes can present an open-shell or a closed-shell structure. Copper phthalocyanine (CuPc) belongs to the first group: the molecule is planar and apolar, with one unpaired electron in the valence shell; this makes it chemically very reactive. Nickel phthalocyanine (NiPc) instead represents a closed-shell system with a strong dipole moment. Among closed-shell MPcs, TiOPc has become of major interest because of its semiconducting, photoconductive<sup>12</sup> and nonlinear optical properties<sup>14</sup>. Moreover, the extremely high chemical and thermal stability of TiOPc allow the formation of high-quality films to be used in electronic and optoelectronic devices<sup>15</sup>. A great peculiarity of TiOPc is that, unlike other MPcs, such as CuPc, it exhibits a non-planar chemical structure: the Ti atom presents a square pyramidal coordination, with the TiO group in axial position, the O atom protruding from the Pc molecular plane, and the four phenyl rings arranged below the Pc molecular plane (see Figure 2); as a consequence, the molecule has a strong dipole moment (about 3.7 D<sup>16</sup>) along the axial direction perpendicular to the Pc molecular plane.



**Figure 2** – Chemical structure of TiOPc: a) top view with labelling of nonequivalent C and N atoms, b) side view.

In the solid state, TiOPc can display various polymorphic phases; each of them possesses a distinct photoconductivity, photodynamic behavior, and charge-transport properties. In literature, there is evidence of four polymorphs, including three monoclinic phases, namely, I, C, and Y,<sup>17,18</sup> and a triclinic phase II<sup>19</sup>. Among them, the Y-form (Y-TiOPc) is an efficient photoconducting material, highly photosensitive to near-infrared (NIR) light, and able to achieve a charge-carrier photogeneration quantum efficiency higher than 90% in high electric fields<sup>20</sup>; this allows it to be largely employed as charge-generation material in laser printers<sup>12</sup>. The triclinic phase II ( $\alpha$ -TiOPc) is instead an excellent *p*-type semiconductor, widely used in solar cells.<sup>21</sup> The structure of phases I and II is shown in Figure 3.

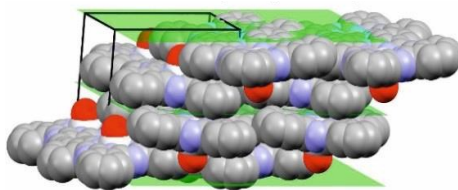


**Figure 3** – Solid phase polymorphism for the TiOPc molecule: a) crystal packing of phase I, a) crystal packing of phase II (adapted from [17]).

On solid/vacuum interfaces, PCs generally adsorb with the molecular plane parallel to the surface, and often form long-range ordered structures at monolayer coverage. Their flat adsorption geometry, combined with their large adsorption energies (generally of several 100 kJ/mol), favors

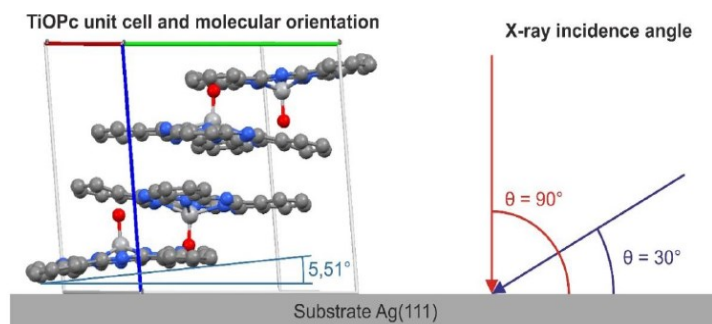
the direct contact of the molecule with the substrate; the resulting monolayers are often stable against thermal desorption, and, at high temperatures, undergo decomposition instead of desorption. One can obtain TiOPc layers adsorbed on different substrates: in the case of graphite, rather smooth films can be formed even at high temperature with a bilayer growth mode; the nonplanar geometry of the molecule generates a strong axial dipole moment which leads to a stacking with the titanyle units alternatingly (*i.e.*, up/down) oriented in adjacent layers<sup>22</sup>. The molecule grows with its backbone parallel to the surface plane and the O atom pointing away from the substrate.<sup>23,24</sup> On the Au(111) surface, TiOPc forms a square lattice of flat-lying molecules, while, on Ag(111), different preparation conditions can lead to various adsorbate structures, including a honeycomb pattern of tilted interlocked pairs, and a hexagonal phase with tilted molecules<sup>25</sup>. Flat-lying TiOPc molecules with the O atom alternatingly oriented can be also observed on Ag(111).<sup>26</sup> However, other studies devoted to TiOPc on Ag(111) also indicated that the O atom points toward the vacuum in the monolayer.<sup>27,28</sup> In the second layer, molecules grow on top of the first one, and the O atom faces toward the surface; this ensures minimal dipole–dipole interactions between layers<sup>29</sup>.

A recent investigation on the growth of TiOPc multilayer films on Ag(111) performed through X-Ray Diffraction (XRD) by the research group of professor G. Witte from the Philipps University of Marburg (Germany), has demonstrated that the TiOPc molecules are epitaxially aligned with respect to the surface, forming an homogeneous film with a regular morphology and a high order of crystallinity, and adopting the phase I polymorph. The (001) molecular planes are oriented nearly parallel to the substrate surface, with an alternating stacking of the titanyle groups forming stable bilayer units (see Figure 4).<sup>30</sup>



**Figure 4** – Molecular arrangement of TiOPc on Ag(111) determined through XRD (adapted from [30]).

Moreover, it has been observed that the TiOPc molecules adopt a reclined orientation with a 5.51° tilt angle of their molecular plane with respect to the crystallographic plane induced by the bulk phase bilayer stacking motive to reduce the lateral repulsions between the molecular dipole moments (see Figure 5).<sup>30</sup>



**Figure 5** – Left side: TiOPc unit cell and molecular orientation on the Ag(111) surface. Right side: definition of X-ray incidence angles measured in the NEXAFS experiment.

Aim of this work is to shed some light on the interaction between the TiOPc molecule and the Ag(111) surface as well as to investigate the dichroic behavior of the molecule, which may be different from that displayed by planar MPs, such as CuPc, due to the bend nature of the TiOPc backbone; to achieve this, NEXAFS spectra at the C,N, and O *K-edge* have been calculated both for the randomly-oriented gas phase molecule and the fixed-in-space molecule, and then compared to the experimental ones, measured at different X-ray incidence angles (30°, 55°, 70°, and 90°) and relative to a multilayer of TiOPc adsorbed on Ag(111). The latter have been recorded at the HESGM Dipole Beamline of the BESSY II Synchrotron of the Helmutz Center (Berlin, Germany) by the research group of professor G. Witte. We point out that the modeling of the Ti *L<sub>II,III</sub>-edge* NEXAFS spectra has not been considered due to the lack of the relative experimental measurements. However, such calculations could be the object of future investigations; the closed-shell nature of the TiOPc molecule and the peculiar electronic configuration of the Ti central ion would allow the employment of the relativistic TDDFT including spin-orbit effects. The nature of the adsorbate-substrate interaction has been investigated by considering the differences in spectral features between the calculated free (*i.e.*, gas phase) molecule and the molecule adsorbed on the Ag(111) surface, with the assumption that, if the adsorbate interacts very weakly with the substrate (physisorption), the corresponding spectral structures are only slightly modified with respect to the free molecule, while if the overlap between the adsorbate and substrate molecular orbitals is more effective (chemisorption), a re-hybridization of the valence levels and consequent changes in the spectra will occur.

## Computational Details

The equilibrium geometry of TiOPc (phase I) extracted from the Cambridge Structural Database (CSD, [17]), has been used as initial guess, and then optimized at the Density Functional Theory (DFT, [31]) level by using the hybrid B3LYP xc potential<sup>32</sup> and the triple  $\zeta$  polarized (TZP) basis set of Slater type orbitals (STOs) extracted from the ADF (Amsterdam Density Functional)

database<sup>33,34</sup>, without imposing any symmetry constraint. Table 1 of the Appendix shows a comparison between experimental<sup>17</sup> and theoretical structural parameters of TiOPc. NEXAFS spectra have been calculated at the C, N, and O *K-edge*, both for the randomly-oriented gas phase molecule and the fixed-in-space molecule with the molecular plane in the xy plane, for which polarized spectra have been computed for all X-ray incidence angles employed in the experimental measurements (30°, 55°, 70°, and 90°) by considering the experimentally observed tilt angle of 5.51° of the molecular xy plane with respect to the X-ray incidence angle<sup>30</sup>. Therefore, the incidence angle considered in these computations corresponds to the difference between the experimental X-ray incidence angle and the tilt angle. Two different exchange correlation (xc) functionals have been tested, namely, the GGA PW86x Perdew (PW86)<sup>35</sup> and the hybrid B3LYP<sup>32</sup>; since the best agreement with the experimental data has been obtained by using the former, only the results obtained with the PW86x xc potential will be discussed in the following Section.

NEXAFS spectra have been calculated through the DFT method within the Transition Potential (TP) scheme<sup>36</sup>, which is well known for its ability to simulate *K-shell* NEXAFS spectra of light atoms<sup>37</sup>. Within this approach, Kohn–Sham (KS) orbitals are determined self-consistently with an orbital occupation scheme obtained by removing half an electron from the excited core orbital, leaving all virtual molecular orbitals (MOs) unoccupied and relaxing all orbitals until self-consistency. Relaxation effects following the core-hole formation are adequately taken into account<sup>38</sup>. As concerns the basis set, we have used the one previously adopted in an analogous work devoted to other MPcs, since it has proven to be accurate enough to deal with this kind of systems<sup>39</sup>. In particular, an even tempered quadruple  $\zeta$  polarized basis set with three polarization and three diffuse functions (designed as ET-QZ3P-3DIFFUSE set in the ADF database) has been employed for the core-excited C, N, and O atoms, while a triple  $\zeta$  polarized basis set with two polarization functions (TZ2P) has been used for the remaining atoms; the core orbitals of the latter have been treated by the Frozen Core (FC) approximation. In particular, a FC TZ2P.1s basis set has been employed for the C, N and O atoms, while a FC TZ2P.3p basis set has been adopted for the Ti atom. The FC approximation consists of treating explicitly only outer level electrons, while the innermost (core) atomic shells are kept frozen. Moreover, this technique ensures the localization of the half core-hole on the excited atom. Within the TP scheme, excitation energies are obtained as the differences between the eigenvalues of the virtual orbital and that of the core orbital calculated with the TP configuration:

$$\Delta E_{i \rightarrow f} = \epsilon_f^{TP} - \epsilon_i^{TP}, \quad (1)$$

while transition intensities are expressed in terms of oscillator strengths,  $f_{i \rightarrow f}$ . For samples in gas phase,  $f_{i \rightarrow f}$  reads:

$$f_{i \rightarrow f} = \frac{2}{3} n_i \Delta E_{i \rightarrow f} |\langle \varphi_f^{TP} | \boldsymbol{\mu} | \varphi_i^{TP} \rangle|^2, \quad (2)$$

involving dipole matrix elements of the electric dipole operator between initial and final TP MOs, where  $n_i$  denotes the occupation number of the core orbital in the ground state. In deriving Eq. (2), one performs an average over all possible orientations of the gas-phase molecule. To compare with experimental NEXAFS intensities recorded for molecules adsorbed on surfaces ( $s$ - and  $p$ -polarized spectra), we consider the molecule fixed in space, and calculate the NEXAFS intensity in the molecular frame (MF) defined such that the Pc plane is parallel to the  $xy$  plane of the MF, while the Ti=O bond defines the MF  $z$ -axis. If  $(\vartheta, \phi)$  are the polar and azimuthal angles of the polarization vector of the incident light in the MF, by performing an average over only the  $\phi$  angle, we obtain the following formula for the oscillator strength for the transition  $i \rightarrow f$ :

$$f_{i \rightarrow f} = 2n_i \Delta E_{i \rightarrow f} \left[ \frac{1}{2} (|\boldsymbol{\mu}_x|^2 + |\boldsymbol{\mu}_y|^2) \sin^2 \vartheta + |\boldsymbol{\mu}_z|^2 \cos^2 \vartheta \right] \quad (3)$$

where  $\vartheta$  is the angle between the  $z$  axis of the molecular frame and the polarization vector of the incident radiation. The quantity in Eq. (3) allows a direct comparison between theoretical and experimental intensities recorded at different X-ray incident angles. In particular, if  $\vartheta = 90^\circ$  ( $s$ -pol), Eq. (3) reduces to:

$$f_{i \rightarrow f} = 2n_i \Delta E_{i \rightarrow f} \left[ \frac{1}{2} (|\boldsymbol{\mu}_x|^2 + |\boldsymbol{\mu}_y|^2) \right], \quad (4)$$

while for  $\vartheta = 0^\circ$  ( $p$ -pol):

$$f_{i \rightarrow f} = 2n_i \Delta E_{i \rightarrow f} |\boldsymbol{\mu}_z|^2 \quad (5)$$

As we will see in the following Section, the analysis of the computed  $s$ -pol and  $p$ -pol intensities of a given transition directly maps the nature  $\pi^*/\sigma^*$  of the virtual MOs implicated in the core-excitation process.

The Ionization Potential (IP) is defined as the negative of the TP eigenvalue related to the initial core orbital,  $IP = -\epsilon_f^{TP}$ . Since the TP approach leads to a less attractive potential and consequently the absolute transition energies are generally too large, the NEXAFS energies have been adjusted by first computing the IPs with the  $\Delta$ KS ( $\Delta$ SCF Kohn-Sham) scheme, allowing a full relaxation of the ionized core hole. The energy of the  $1s^{-1}$  ionic state has been obtained through a KS unrestricted calculation. Hence, the TP excitation energies (Eq. 1) have been shifted with respect to the  $\Delta$ IP value by an amount given by the energy difference  $\epsilon_{1s}^{TP} - \Delta$ KS. A separate computation of the excitation spectrum for each nonequivalent C and N atomic site has been performed, and the total spectrum has been obtained by summing up the partial contributions weighted by the number of equivalent C or N atoms. In particular, eight equivalent C4 sites (“pyrrolic” C atoms) and eight equivalent C1, C2 and C3 sites (“phenyl” C atoms) have been considered in C1s NEXAFS

calculations, while four equivalent N atoms of the pyrrole rings bonded to the metal atom ( $N_{\text{pyr}}$ ) and four equivalent N atoms in the *meso* positions of the macrocyclic ring ( $N_{\text{aza}}$ ) have been taken into account in N1s calculations (see Figure 2 for the corresponding labeling). Preliminary calculations have been performed to check the equivalence of the C1, C2, C3, and C4 sites to C atoms occupying the same relative positions in the phenyl and pyrrolic rings they belong to, as well as of all four  $N_{\text{aza}}$  and  $N_{\text{pyr}}$  sites. These tests are reported in Figure 2 of the Appendix, together with the chemical structure of the molecule showing the labeling of all C and N sites considered in these computations.

Line spectra have been convoluted with Gaussian functions of appropriate full-width at half maximum (FWHM) value (equal to 0.5 eV for C1s spectra, and to 0.7 eV for both N1s and O1s spectra). In order to get an easier comparison with the experimental data, the theoretical C1s and O1s line-shapes have been shifted by -0.2 eV, while a rigid shift of +0.45 eV has been applied to the computed N1s spectra.

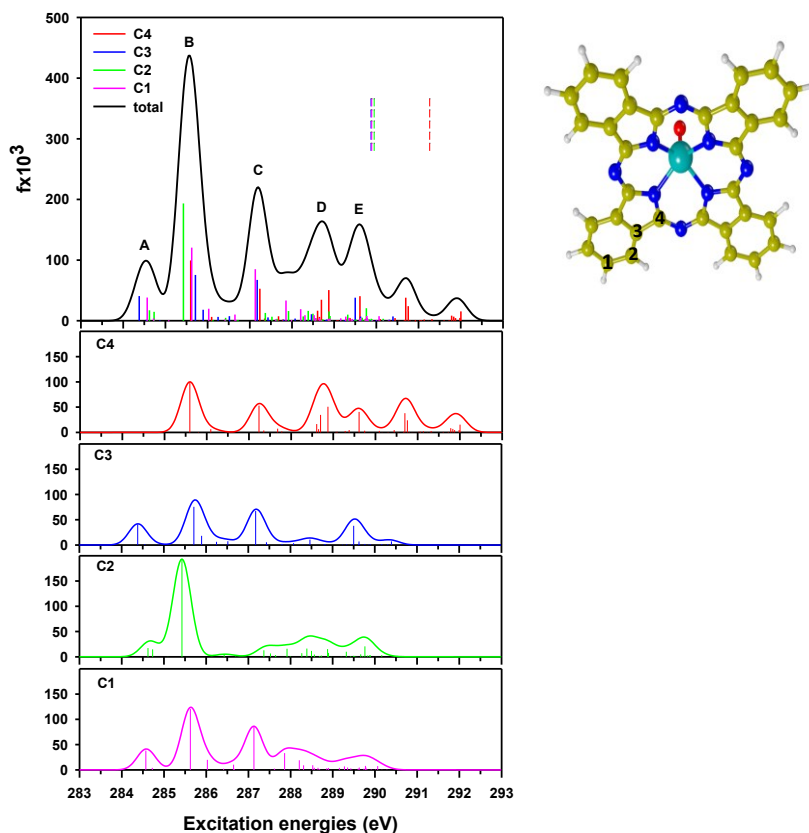
## Results and Discussion

In the following subsections, the TP-DFT results will be discussed in order to provide an assignment of the experimental NEXAFS spectral features. The analysis will be limited to the below-edge region, since transitions to virtual MOs in the electronic continuum cannot be accurately described through the employed computational protocol. In the first subsection, NEXAFS spectra of the randomly-oriented gas phase molecule will be considered, while, in the second one, polarized NEXAFS spectra will be presented.

### C1s, N1s and O1s NEXAFS spectra of randomly-oriented gas phase TiOPc

The theoretical C1s NEXAFS spectrum of randomly-oriented gas phase TiOPc. is displayed in Figure 6, together with the partial contributions of all nonequivalent  $C_i$  sites. The chemical structure of the molecule reporting the labeling of the nonequivalent  $C_i$  sites is also shown. The deconvolution of the total spectral profile into site-resolved components facilitates the assignments of the spectral features to specific portions of the molecule. The assignment of the spectral features is reported in Table 1. Since the virtual MO set is modified by the creation of the core-hole, the reported energies, intensities, and atomic contributions are specific for each nonequivalent excitation site. The MOs involved in the most intense transitions have been identified, and their composition analyzed in terms of Mulliken population of atomic orbitals (AOs) that are centered on

each specific atom; the latter has been included in Table 2 of the Appendix, where only the main atomic contributions, which are useful to the attribution of the spectral features, are reported.



**Figure 6** – Left side: C1s NEXAFS spectrum of randomly oriented gas-phase TiOPc (upper panel). Vertical colored dashed bars:  $\Delta$ KS IPs (C1=289.89 eV, C2=289.95 eV, C3=289.87 eV, C4=291.26 eV). Site-resolved  $C_i$  contributions are presented as colored spectra in the lower panels. Right side: chemical structure of TiOPc with the labeling of nonequivalent C atoms.

**Table 1** – Peak assignments of the C1s NEXAFS spectrum of randomly-oriented gas phase TiOPc. Only transitions with  $fx10^3 \geq 16.0$  are reported.

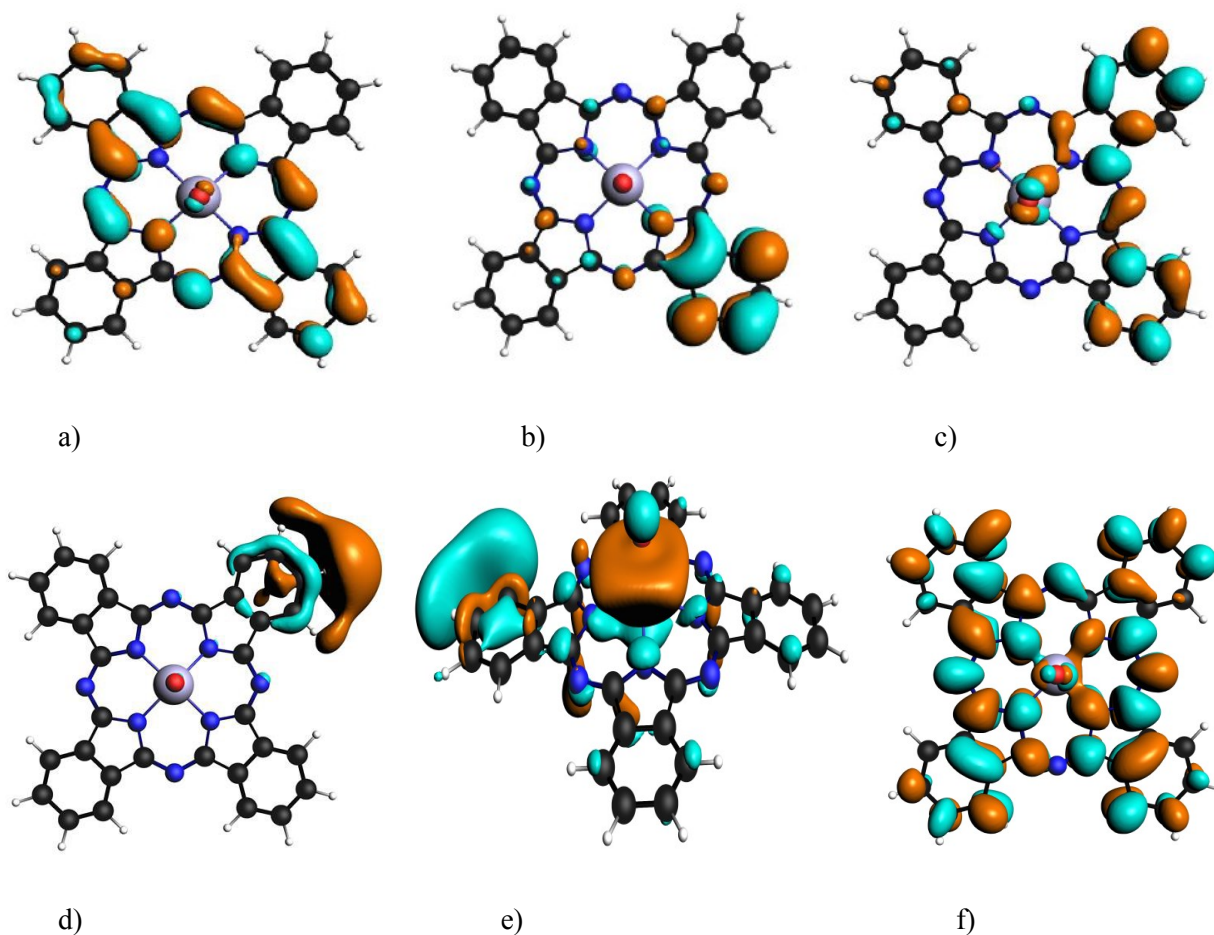
Peak	Site	E (eV)	$fx10^3$	Assignment
A	C3	284.38	40.8	$\pi^*(C=C)+\pi^*(C-N)$ (LUMO)
	C1	284.57	38.2	
	C2	284.62	17.3	
B	C2	285.43	193.0	$\pi^*(C=C)+\pi^*(C-N)$ (LUMO)
	C4	285.60	99.4	
	C1	285.62	121.0	$\pi^*(C=C)+\pi^*(C-N)$
	C3	285.71	75.6	
	C3	285.90	18.2	
	C1	286.03	19.9	
C	C1	287.13	85.1	$\pi^*(C=C)+\pi^*(C-N)$
	C3	287.18	67.4	$\pi^*(C=C)+3d Ti$
	C4	287.24	52.8	$\pi^*(C=C)+\pi^*(C-N)+3d Ti$
D	C1	287.86	33.4	$\sigma^*(C-H)$
	C2	287.92	16.1	



	C1	288.21	19.1	$\sigma^*(\text{C-H})+\sigma^*(\text{Ti=O})$
	C2	288.39	16.4	
	C4	288.61	16.5	$\pi^*(\text{Ti=O})+\pi^*(\text{C=C})$ $\pi^*(\text{C=C})+3d \text{ Ti}$
	C4	288.70	34.7	
	C4	288.87	50.6	
<b>E</b>	C3	289.50	38.0	$\pi^*(\text{C=C})+\pi^*(\text{C-N})+np \text{ C}$
	C4	289.61	40.5	$\pi^*(\text{C=C})+\pi^*(\text{C-N})+3d \text{ Ti}$
	C2	289.76	20.7	$\pi^*(\text{C=C})+\pi^*(\text{C-N})+ns,np \text{ C}$

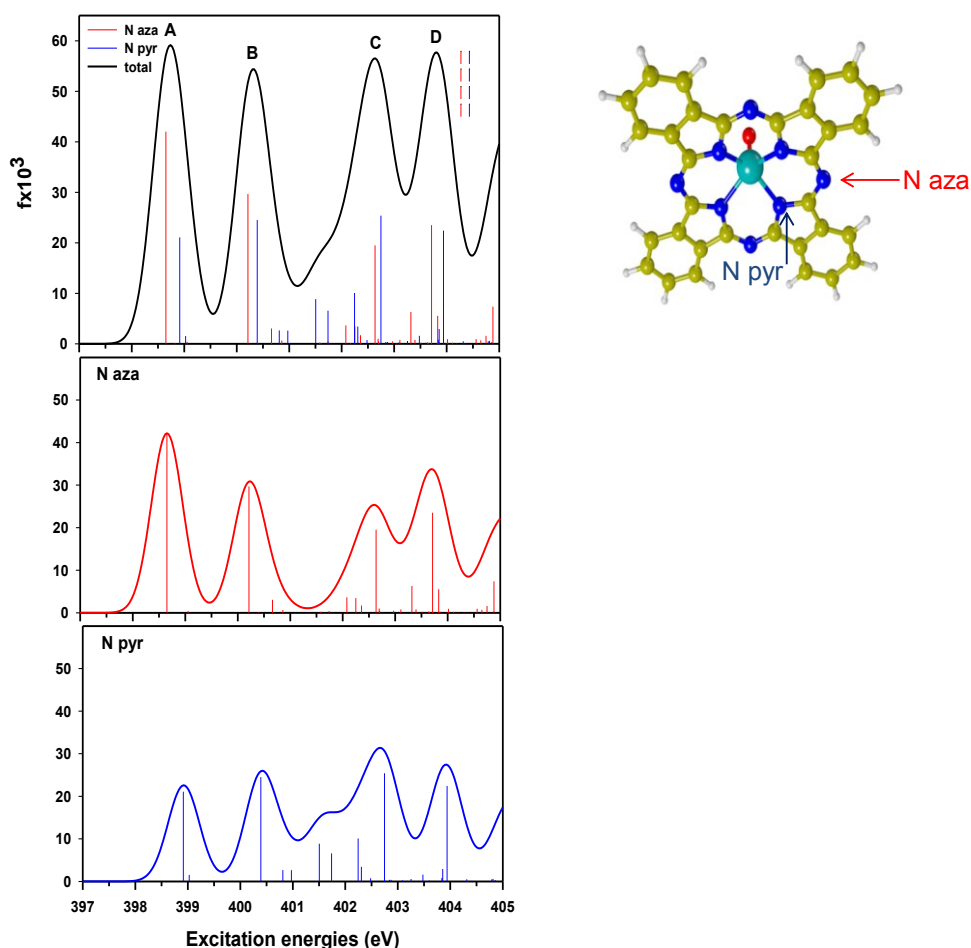
The total calculated C1s NEXAFS spectrum is dominated by an intense double-peaked feature (labeled A and B in Figure 6). Peak A mainly derives from the LUMO ( $1\pi^*$ ) transitions of the C atoms of the phenyl rings; among them, the transition relative to the C2 site displays the lowest intensity, in line with the small C 2p AO participation to this MO (see Table 2 of the Appendix). The LUMO exhibits a strong C  $2p_\pi$  character of the “pyrrolic” rings and, to a minor extent, of the outer benzene rings, as it can be seen from Figure 7a. The LUMO transition from the “pyrrolic” C atom (C4) falls in the energy region of peak B, shifted to higher energy by about 1.0 eV compared to the LUMO transitions relative to the “phenyl” C sites; this can be rationalized by considering the inductive effect of the neighboring N atoms on the “pyrrolic” C atoms. Moreover, the trend of the LUMO transitions resembles that followed by the calculated IPs, and is therefore ascribable to initial state effects. We point out that the IP trend (C1-C3 and C4) follows that already observed in polycyclic aromatic hydrocarbons, according to which higher IPs are predicted for C atoms not bonded to any H atom compared to external carbons bonded to H atoms<sup>40</sup>. Peak B, which is the most intense feature of the total spectrum, is also contributed by transitions from the C atoms of the phenyl rings towards MOs which present the same character of the LUMO. In particular, the most intense transition contributing to this peak is relative to C2, and occurs at 285.43 eV (see Table 1); the high intensity of this transition reflects the significant  $2p_z$  AO participation of the C atom carrying the core-hole (around 23%), as one can observe from the plot of the corresponding MO (Figure 7b). The second most intense transition of peak B originates from the C1 site (at 285.62 eV, see Table 1). Peak C derives its intensity from the C1, C3 and C4 transitions toward MOs with a main  $\pi^*(\text{C=C})$  character and minor  $\pi^*(\text{C-N})$  and Ti 3d contributions; the plot of a representative MO is shown in Figure 7c. The C1 and C2 transitions to  $\sigma^*(\text{C-H})$  MOs are responsible of the overlap between peaks C and D; the plot of the final MO relative to the C1 transition is displayed in Figure 7d. The lower-energy shoulder of peak D arises from the C1 and C2 excitations toward final MOs with mixing  $\sigma^*(\text{C-H})$  and  $\sigma^*(\text{Ti=O})$  contributions, as one can observe from Figure 7e. The most intense transitions of peak D are relative to C4, and are toward virtual MOs which present a  $\pi^*(\text{C=C})$  character with consistent metal 3d contributions; the less intense transition from C4 is

toward a MO with a main  $\pi^*(\text{Ti}=\text{O})$  character and minor  $\pi^*(\text{C}=\text{C})$  contributions. Just below the ionization threshold, peak E is visible: it is ascribed to transitions originating from C2, C3 and C4 sites toward MOs with a main  $\pi^*(\text{C}=\text{C})$  character and minor  $\pi^*(\text{C}-\text{N})$  contributions; lower weights of C ( $ns, np$ ) AOs and Ti 3d contributions are also visible (see Figure 7f).



**Figure 7** – Plots of representative final MOs involved in the most intense C1s NEXAFS transitions: a) LUMO (transition at 284.57 eV), b)  $\pi^*(\text{C}=\text{C}) + \pi^*(\text{C}-\text{N})$  (transition at 285.43 eV), c)  $\pi^*(\text{C}=\text{C}) + 3d \text{ Ti}$  (transition at 287.18 eV), d)  $\sigma^*(\text{C}-\text{H})$  (transition at 287.86 eV), e)  $\sigma^*(\text{C}-\text{H}) + \sigma^*(\text{Ti}=\text{O})$  (transition at 288.39 eV), f)  $\pi^*(\text{C}=\text{C}) + \pi^*(\text{C}-\text{N}) + ns, np \text{ C}$  (transition at 289.76 eV).

The total calculated N1s NEXAFS spectrum of the randomly-oriented gas phase molecule as well as the partial  $N_i$  contributions are displayed in Figure 8, together with the chemical structure of the molecule reporting the labeling of the nonequivalent  $N_i$  sites. The assignments of the most intense transitions of the spectrum are listed in Table 2. Table 3 of the Appendix also includes the composition of the implicated MOs in terms of Mulliken population of AOs centered on each specific atom.



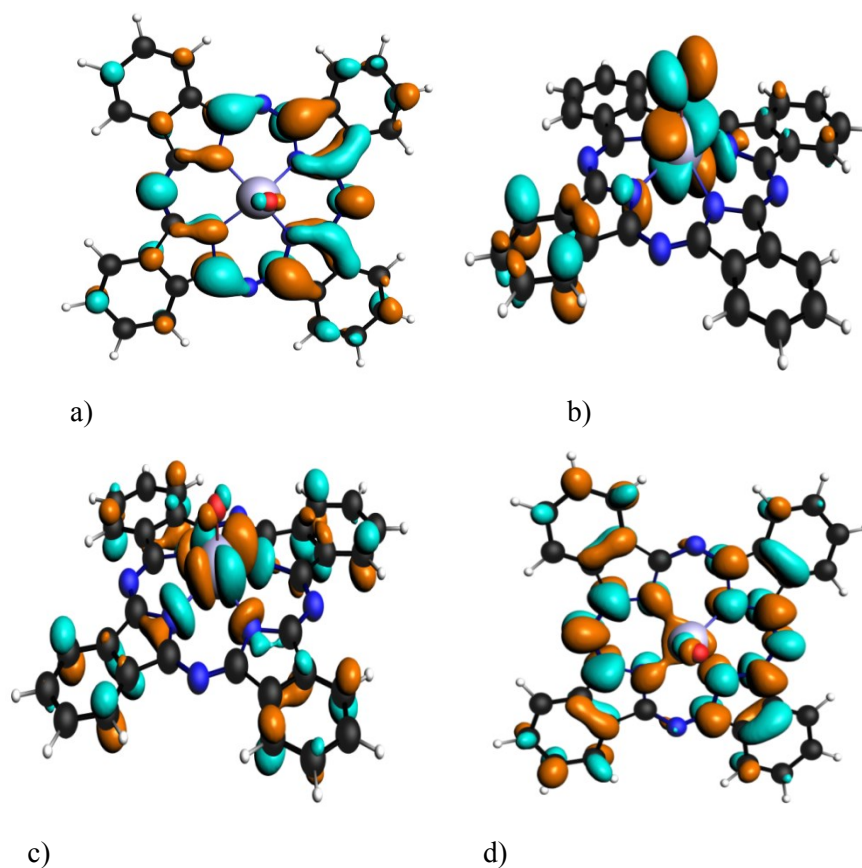
**Figure 8** – Left side: N1s NEXAFS spectrum of randomly-oriented gas phase TiOPc (upper panel). Vertical colored dashed bars:  $\Delta$ KS IPs ( $N_{\text{aza}} = 404.27$  eV,  $N_{\text{pyr}} = 404.43$  eV). Partial  $N_i$  contributions are displayed as colored spectra in the bottom panels. Right side: chemical structure of TiOPc with the labeling of nonequivalent N atoms.

**Table 2** – Peak assignments of the N1s NEXAFS spectrum of randomly-oriented gas phase TiOPc. Only transitions with  $\text{fx}10^3 \geq 5.00$  are reported.

Peak	Site	E (eV)	$\text{fx}10^3$	Assignment
A	$N_{\text{aza}}$	398.65	42.00	$\pi^*(\text{C}=\text{C}) + \pi^*(\text{C}-\text{N})$ (LUMO)
	$N_{\text{pyr}}$	398.91	21.10	
B	$N_{\text{aza}}$	400.21	29.70	$\pi^*(\text{C}=\text{C}) + \pi^*(\text{C}-\text{N})$
	$N_{\text{pyr}}$	400.39	24.50	

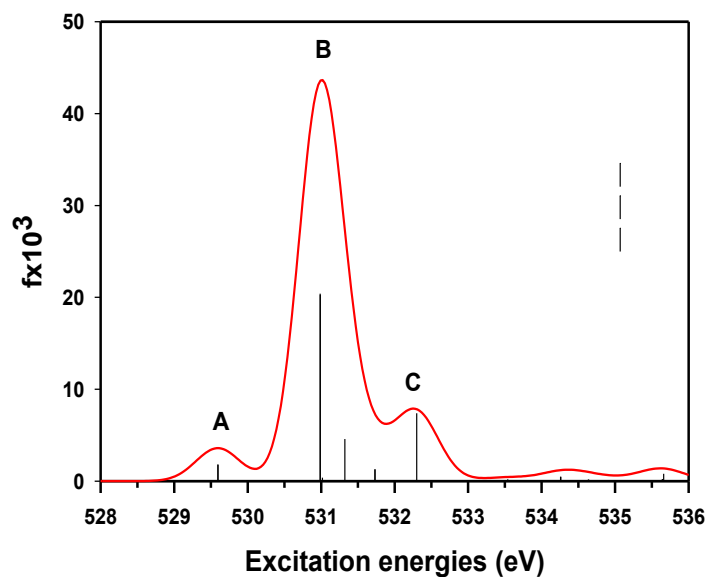
<b>C</b>	N <sub>pyr</sub>	401.51	8.87	$\pi^*(\text{Ti}=\text{O})+\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})$
	N <sub>pyr</sub>	401.74	6.58	$\pi^*(\text{C}=\text{C})+3d \text{ Ti}+\sigma^*(\text{C}-\text{N})$
	N <sub>pyr</sub>	402.24	10.10	$\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})$
	N <sub>aza</sub>	402.63	19.50	
	N <sub>pyr</sub>	402.75	25.40	$\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})+3d \text{ Ti}$
<b>D</b>	N <sub>aza</sub>	403.32	6.31	Rydberg (N)+ $\pi^*(\text{C}=\text{C})$
	N <sub>aza</sub>	403.71	23.50	$\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})+3d, np \text{ N}$
	N <sub>aza</sub>	403.83	5.55	Rydberg (N)+ $\pi^*(\text{C}=\text{C})$
	N <sub>pyr</sub>	403.94	22.40	$\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})+3d \text{ N}$

The total calculated N1s NEXAFS spectrum appears less structured than the C1s one due to the lower number of nonequivalent core-hole sites: it is dominated by a low-energy peak (A) which is contributed by the N(1s)→LUMO transitions from the two non-equivalent N sites. The energy splitting between the two transitions to the LUMO (equal to 0.26 eV) is larger than the splitting between the two N core-holes (equal to 0.16 eV); this accounts for the different relaxation of the LUMO upon the core-hole formation on the two non-equivalent sites. The significant oscillator strength calculated for the LUMO transitions reflects the high weight of N<sub>aza</sub> and N<sub>pyr</sub> 2p<sub>π</sub> components into the LUMO orbital, which exhibits a C 2p<sub>π</sub> character of the pyrrolic rings, and, to a minor extent, of the outer benzene rings (see Figure 9a). Peak B is contributed by two intense transitions from both N<sub>aza</sub> and N<sub>pyr</sub> sites to final MOs which essentially present the same character of the LUMO. The strongest transitions contributing to peak C are toward MOs which exhibit a strong  $\pi^*(\text{C}=\text{C})$  character and minor  $\pi^*(\text{C}-\text{N})$  contributions, with the exception of the transition at 401.51 eV which involves a  $\pi^*(\text{Ti}=\text{O})$  MO with some  $\pi^*(\text{C}-\text{N})$  and  $\pi^*(\text{C}=\text{C})$  contributions (Figure 9b). The transition at 401.74 eV is toward a MO which presents a main Ti 3d character and minor  $\pi^*(\text{C}=\text{C})$  contributions and a less pronounced  $\sigma^*(\text{C}-\text{N})$  character (N 2p<sub>x,y</sub> contributions around 4%, see Table 3 of the Appendix), as it is possible to observe from the relative plot (see Figure 9c). The two most intense transitions responsible of peak D are from both N sites, and involve virtual MOs with a strong C 2p<sub>π</sub> and N 2p<sub>π</sub> character and lower weights of N (np,3d) AOs (see Figure 9d). In the same spectral region, transitions of lower intensity to MOs with a strong Rydberg character and minor  $\pi^*(\text{C}=\text{C})$  contributions are also visible.



**Figure 9** – Plots of representative final MOs involved in the most intense N1s NEXAFS transitions: a) LUMO (transition at 398.65 eV), b)  $\pi^*(\text{Ti}=\text{O})+\pi^*(\text{C}=\text{C})$  (transition at 401.51 eV), c)  $\pi^*(\text{C}=\text{C})+3d$  Ti (transition at 401.74 eV), d)  $\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})+3d, np$  N (transition at 403.71 eV).

In Figure 10, the total calculated O1s NEXAFS spectrum of the randomly-oriented gas phase molecule is displayed, and the assignments of the absorption bands are listed in Table 3. Table 4 of the Appendix also reports the composition of the implicated MOs in terms of Mulliken population of AOs centered on each specific atom.

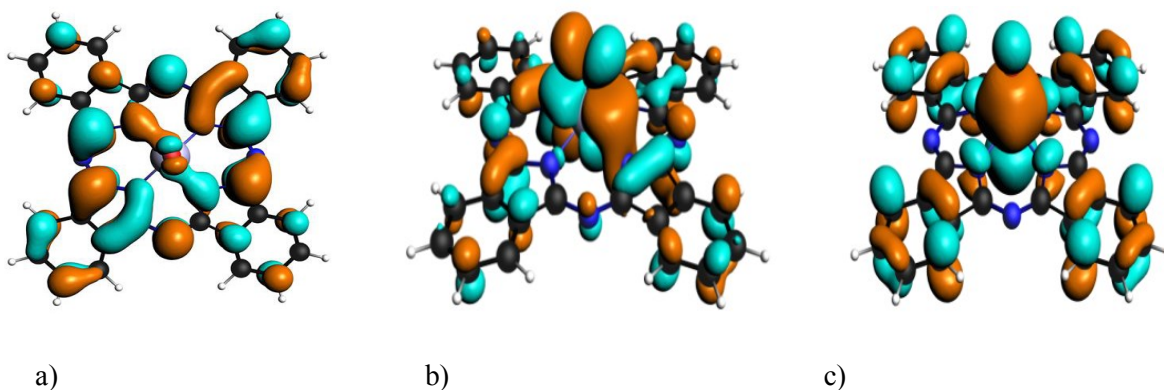


**Figure 10** – O1s NEXAFS spectrum of randomly-oriented gas phase TiOPc. Vertical dashed bar:  $\Delta$ KS IP (equal to 535.07 eV).

**Table 3** – Peak assignments of the O1s NEXAFS spectrum of randomly-oriented gas phase TiOPc. Only transitions with  $fx10^3 \geq 1.00$  are reported.

Peak	E (eV)	$fx10^3$	Assignment
A	529.59	1.81	$\pi^*(C=C) + \pi^*(C-N)$ (LUMO+1)
B	530.99	20.4	$\pi^*(Ti=O)$
	531.32	4.58	$\sigma^*(Ti=O) + \pi^*(C=C)$
	531.73	1.29	$\pi^*(Ti=O) + \pi^*(C=C)$
C	532.30	7.39	$\sigma^*(Ti=O) + \pi^*(C=C)$

At lower excitation energies, the O1s NEXAFS spectrum is characterized by a weak peak (A) which arises from a transition toward the LUMO+1 orbital. From the plot of the MO reported in Figure 11a, it clearly exhibits a strong C  $2p_\pi$  character of the pyrrolic rings and of the outer benzene rings. The O(1s)  $\rightarrow$  LUMO transition has a negligible intensity (not reported in Table 3), since the O 2p participation to this MO is very small (see Table 4 of the Appendix). Peak B comprises a manifold of transitions where the final MOs have either  $\pi^*(Ti=O)$  or  $\sigma^*(Ti=O)$  together with a  $\pi^*(C=C)$  character (see Figure 11b and c, respectively). Peak C instead arises from a relatively intense transition which involves a  $\sigma^*(Ti=O)$  MO with some  $\pi^*(C=C)$  contributions from the outer phenyl rings.



**Figure 11** – Plots of representative final MOs involved in the most intense O1s transitions: a)  $\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})$  (transition at 529.59 eV), b)  $\pi^*(\text{Ti}=\text{O})$  (transition at 530.99 eV), c)  $\sigma^*(\text{Ti}=\text{O})+\pi^*(\text{C}=\text{C})$  (transition at 531.32 eV).

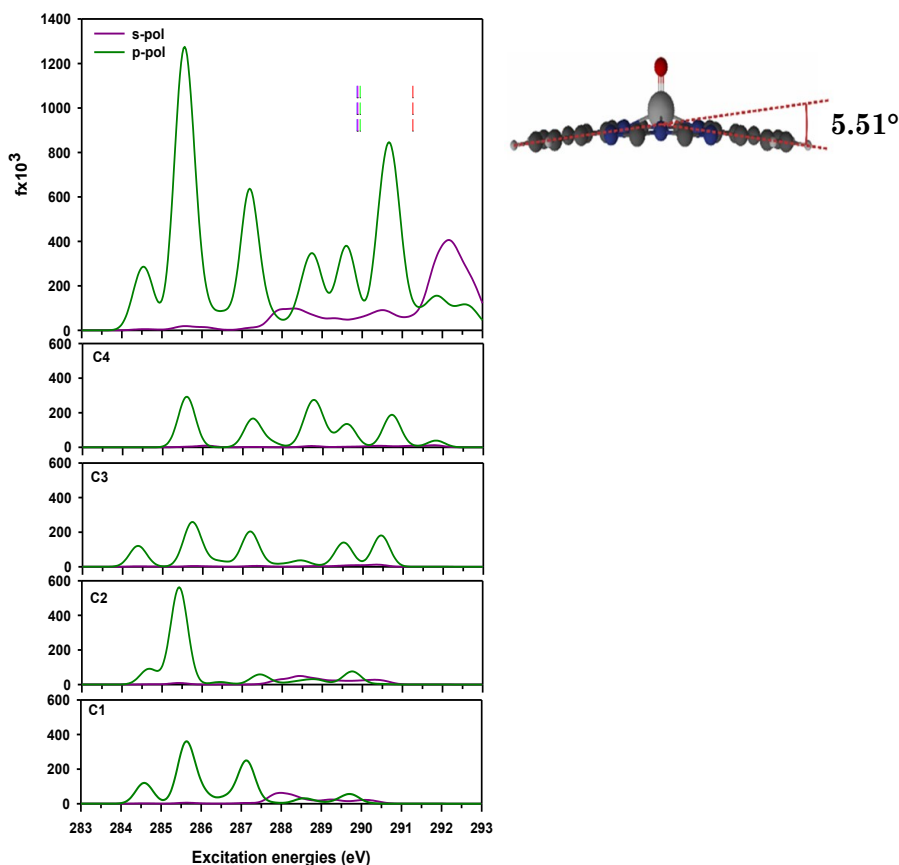
### C,N,O 1s polarized NEXAFS spectra of TiOPc

In order to validate the assignment of the total calculated C,N,O 1s NEXAFS spectra of the gas-phase molecule, we investigated the angular dependence of the intensity of the main transitions. This has been done through the simulation of the *s*- and *p*-polarized NEXAFS spectra. Furthermore, to facilitate the analysis, we calculated the value of the angle ( $\vartheta$ ) of the transition dipole moment (TDM) vector with respect to the molecular *z* axis for the most intense transitions by means of the following formula:

$$\cos \vartheta = \sqrt{\frac{f_{p-pol}}{2f_{s-pol} + f_{p-pol}}} \quad (6)$$

where *f* denotes the oscillator strength calculated for the *s*- and *p*-polarizations. The values of  $\vartheta$  have been collected in Tables 4, 5, and 6, for C1s, N1s, and O1s NEXAFS spectra, respectively.

In Figure 12, the angle-resolved (*s*- and *p*-pol) C1s NEXAFS spectrum is displayed together with the site-resolved *s*- and *p*-pol partial contributions of each nonequivalent C atom. In the analysis, it is necessary to keep in mind that the molecule lies in the *xy* plane of the molecular frame, with the Ti=O bond directed along the *z* axis, and a tilt of 5.51° of the molecular *xy* plane with respect to the X-ray incidence angle (see right side of Figure 12).



**Figure 12** – Left side: polarized (*s*- and *p*-pol) C1s NEXAFS spectrum of gas-phase TiOPc (upper panel). Vertical colored dashed bars:  $\Delta$ KS IPs. Partial  $C_i$  contributions are presented in the lower panels. Right side: orientation of the TiOPc molecule with respect to the molecular frame.

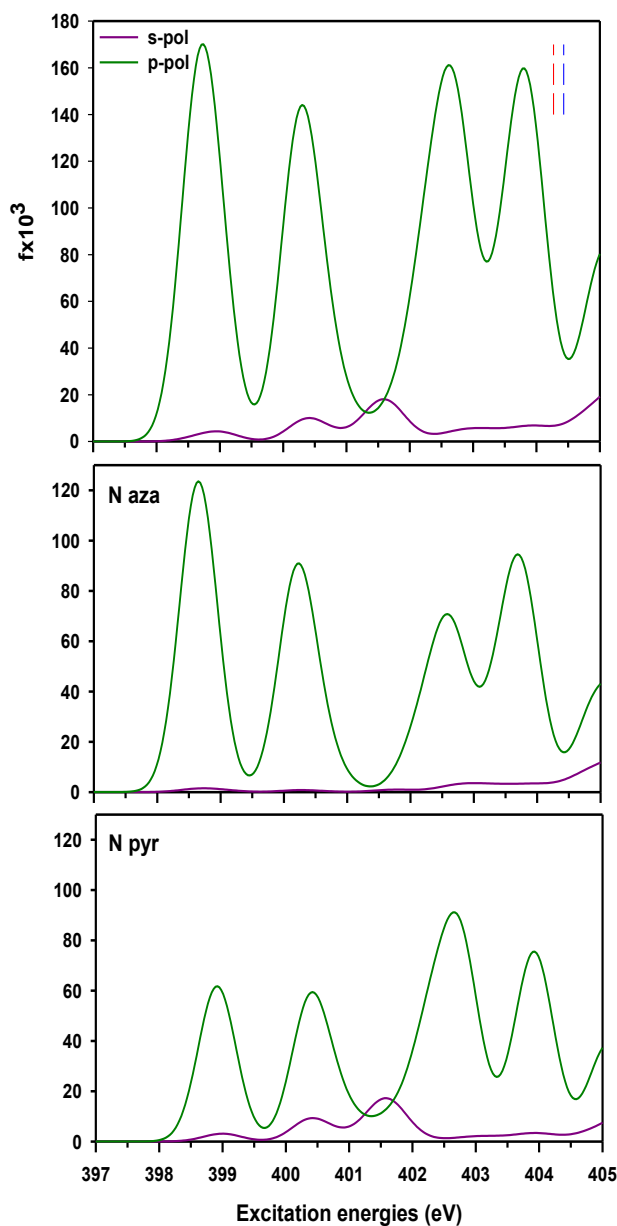
From an analysis of Figure 12, one can observe that the intensity trend reflects the  $\pi$  symmetry of the final MOs which contribute to the main spectral features. Moreover, in accordance with the peak assignment of the C1s NEXAFS spectrum (see Table 1), the first three spectral features are contributed by transitions toward  $\pi^*$  MOs, whereas the final MOs implicated in transitions occurring at around 288 eV display a  $\sigma^*$  character. By looking at the site-resolved partial contributions, it is possible to notice that the *s*-pol contribution is not negligible only in the C1 and C2 partial contributions; this can be rationalized by considering that, unlike C3 and C4, C1 and C2 atoms are both bonded to a H atom, whose *s* orbitals can participate to the formation of  $\sigma^*$  MOs. This is also apparent in the analysis of the  $\vartheta$  angle of the TDM vector, which is almost perpendicular to the *z* axis for  $\sigma^*$  transitions, and nearly parallel to it in the case of  $\pi^*$  transitions (see Table 4); this is in line with the orientation of the molecule, according to which  $\pi^*$  MOs extend mainly along the *z* axis, while  $\sigma^*$  MOs lie parallel to the *xy* plane.



**Table 4** – Values (in °) of the  $\vartheta$  angle that the TDM vector forms with the z axis of the molecular frame of reference for the most intense transitions of the C1s NEXAFS spectrum of gas-phase TiOPc.

Peak	Site	E (eV)	$f_{s-pol} \times 10^3$	$f_{p-pol} \times 10^3$	$\vartheta$ (°)
<b>A</b>	C3	284.38	2.32	118	11.22
	C1	284.57	1.83	111	10.29
	C2	284.62	0.84	50.3	10.36
<b>B</b>	C2	285.43	8.92	562	10.10
	C4	285.60	3.21	292	11.91
	C1	285.62	5.53	351	14.20
	C3	285.71	2.95	221	13.09
	C3	285.90	1.67	51.2	14.33
	C1	286.03	1.05	57.6	10.81
	<b>C</b>	C1	287.13	4.20	2.47
C3		287.18	2.44	197	8.94
C4		287.24	1.82	155	8.71
<b>D</b>	C1	287.86	49.6	0.82	84.81
	C2	287.92	24.1	$1.92 \times 10^{-2}$	88.86
	C1	288.21	28.2	0.92	82.72
	C2	288.39	24.2	0.76	82.86
	C4	288.61	2.14	45.3	17.09
	C4	288.70	3.80	96.4	15.68
	C4	288.87	2.14	148	9.65
<b>E</b>	C3	289.50	1.02	112	7.69
	C4	289.61	1.30	119	8.41
	C2	289.76	2.67	56.7	17.06

In Figure 13, the polarized (*s*- and *p*-pol) N1s NEXAFS spectrum is shown together with the site-resolved *s*- and *p*-pol partial contributions of the two nonequivalent N sites.



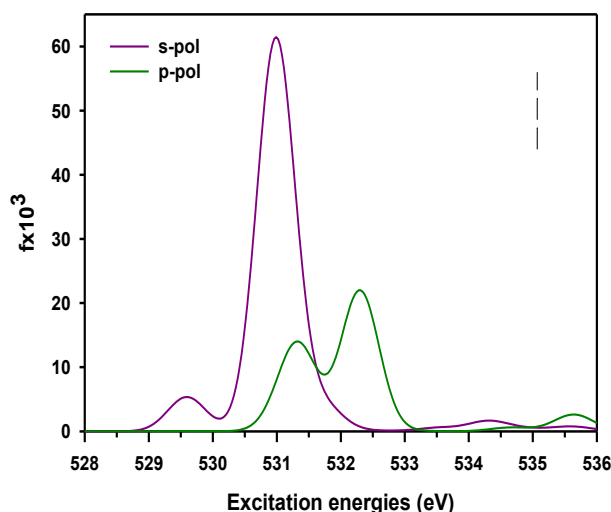
**Figure 13** – Polarized (*s*- and *p*-pol) N1s NEXAFS spectrum of gas-phase TiOPc (upper panel). Vertical colored dashed bars:  $\Delta$ KS IPs. Partial  $N_i$  contributions are presented in the lower panels.

As observed in the C1s case, the results displayed in Figure 13 are consistent with the  $\pi^*$  character of the final MOs implicated in the most intense transitions (see Table 2). A noticeable contribution of the *s*-pol contribution from the  $N_{\text{pyr}}$  core-hole site is found in the energy region between 400 and 402 eV; this is consistent with the  $\sigma^*(\text{C-N})$  contributions to the final MOs, and also reflects the higher value of  $\vartheta$  calculated for these transitions (see Table 5).

**Table 5** – Values (in °) of the  $\vartheta$  angle that the TDM vector forms with the z axis of the molecular frame of reference for the most intense transitions of the N1s NEXAFS spectrum of gas-phase TiOPc.

Peak	Site	E (eV)	$f_{s-pol} \times 10^3$	$f_{p-pol} \times 10^3$	$\vartheta$ (°)
<b>A</b>	N <sub>aza</sub>	398.65	1.24	124	8.05
	N <sub>pyr</sub>	398.91	0.77	61.7	8.98
<b>B</b>	N <sub>aza</sub>	400.21	0.61	87.8	6.72
	N <sub>pyr</sub>	400.39	8.75	56.0	29.2
<b>C</b>	N <sub>pyr</sub>	401.51	11.6	3.34	69.2
	N <sub>pyr</sub>	401.74	6.63	6.48	55.0
	N <sub>pyr</sub>	402.24	0.38	29.5	9.12
	N <sub>aza</sub>	402.63	0.54	57.4	7.81
	N <sub>pyr</sub>	402.75	0.30	75.6	5.09
<b>D</b>	N <sub>aza</sub>	403.32	$7.48 \times 10^{-2}$	18.8	5.10
	N <sub>aza</sub>	403.71	0.28	70.0	5.11
	N <sub>aza</sub>	403.83	$7.34 \times 10^{-2}$	16.5	5.39
	N <sub>pyr</sub>	403.94	1.97	63.4	14.0

Figure 14 reports the polarized (*s*- and *p*-pol) O1s NEXAFS spectrum.



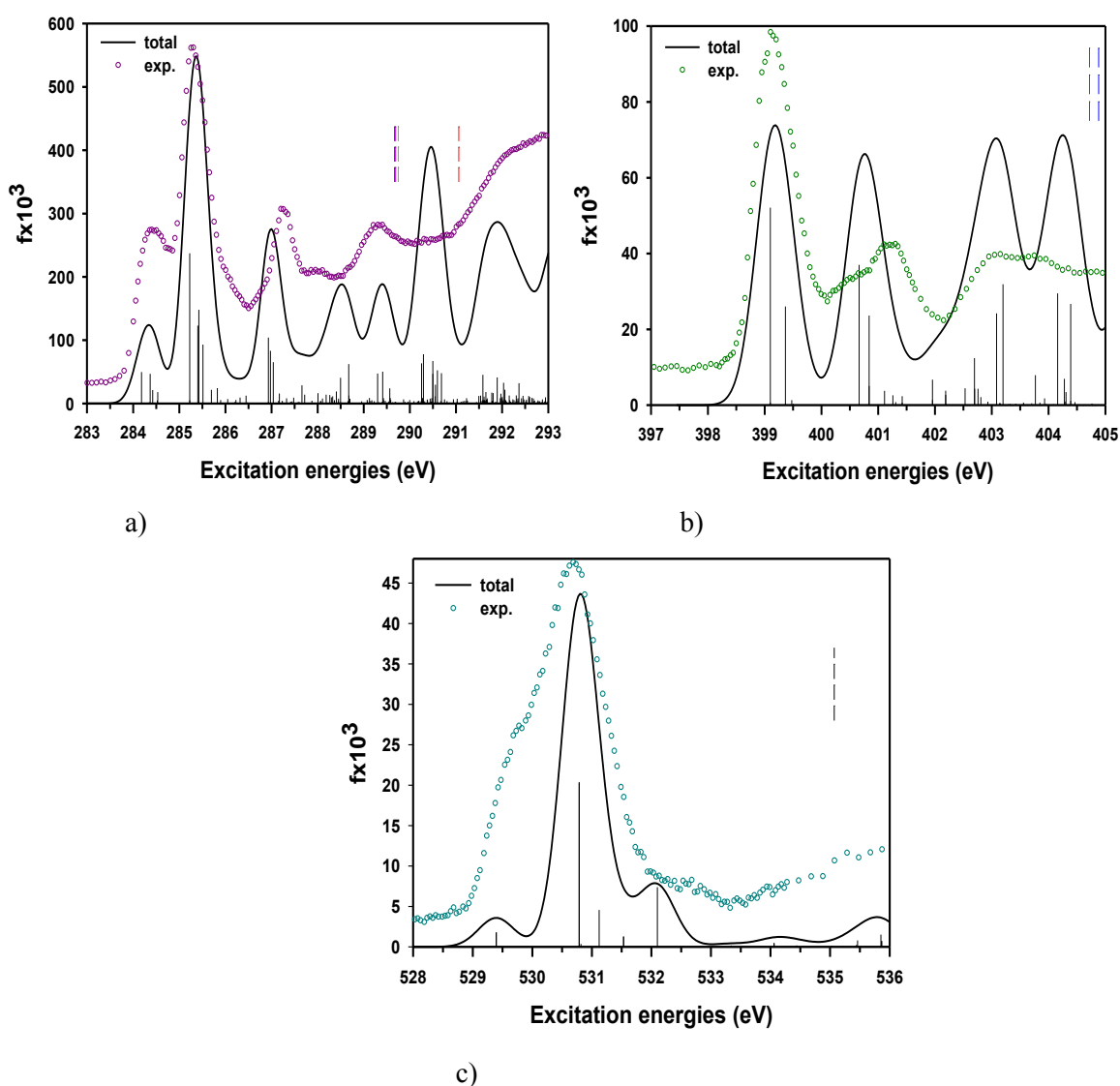
**Figure 14** – Polarized (*s*- and *p*-pol) O1s NEXAFS spectrum of gas-phase TiOPc. Vertical dashed bar:  $\Delta$ KS IP.

An inspection of Figure 14 reveals that, in accordance with the peak assignment of the O1s NEXAFS spectrum (see Table 3), the *s*- and *p*-pol contributions are related to the  $\pi^*$  and  $\sigma^*$  characters of the final MOs implicated in the main transitions, respectively. This is also confirmed by the calculated values of the  $\vartheta$  angle which show that the TDM vector is nearly perpendicular to the z axis in the case of  $\pi^*$  transitions, while it is oriented along the z axis for  $\sigma^*$  transitions (see Table 6). Indeed, the orientation of the molecule implies that  $\pi^*(\text{Ti}=\text{O})$  MOs extend mainly parallel to the xy plane, while  $\sigma^*(\text{Ti}=\text{O})$  MOs along the z axis.

**Table 6** – Values (in °) of the  $\vartheta$  angle that the TDM vector forms with the z axis of the molecular frame of reference for the most intense transitions of the O1s NEXAFS spectrum of gas-phase TiOPc.

Peak	E (eV)	$f_{s-pol} \times 10^3$	$f_{p-pol} \times 10^3$	$\vartheta$ (°)
A	529.59	2.71	$1.03 \times 10^{-3}$	89.2
B	530.99	30.4	0.40	85.4
	531.32	$7.02 \times 10^{-2}$	13.6	5.80
C	531.73	1.92	$1.65 \times 10^{-2}$	86.2
	532.30	0.11	21.9	5.72

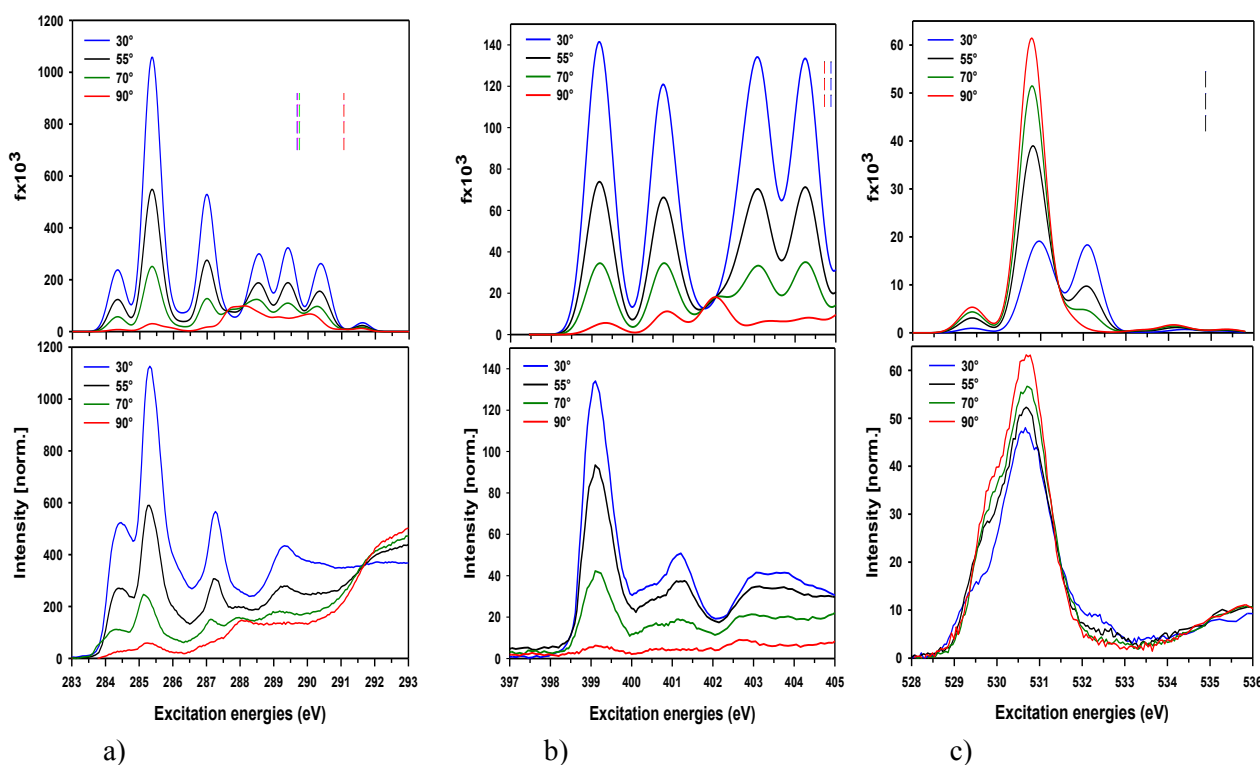
The C,N,O 1s NEXAFS spectra calculated at a 55° X-ray incidence angle are reported in Figure 15, together with the corresponding experimental data.



**Figure 15** – NEXAFS spectra of TiOPc at a 55° X-Ray incidence angle: a) C1s, b) N1s, c) O1s. The experimental spectra (colored circles) are shown together with the total theoretical results (solid black lines). Vertical dashed bars:  $\Delta$ KS IPs. The calculated peaks have been shifted by -0.2 eV (C1s and O1s) and +0.45 eV (N1s) for a better agreement with the experimental data.

A reasonably good agreement between experiment and theory is observed especially in the C1s case, as concerns both the energy positions and the intensity distribution of the first three spectral features. In the region close to the ionization thresholds, the comparison with the theory is more qualitative, as the experimental spectrum presents a broader shape which is split into three peaks in the calculated one. The agreement between theory and experiment is only qualitative in the case of the N1s spectrum, especially as concerns the intensity of the second peak; this could be attributed to the vibrational effects and/or intermolecular interactions neglected in the computational protocol. Finally, it is worth noting that, in the O1s spectrum, the experimental profile is not able to resolve fine structures, in particular as concerns the first and the third theoretical peak, which are present as low and high energy shoulders of the intense second peak.

Figure 16 compares the C,N,O 1s NEXAFS spectra taken at different X-ray incidence angles: the theoretical spectra are reported in the upper part of each panel, while the experimental ones are displayed in the lower part.



**Figure 16** – Polarized NEXAFS spectra of TiOPc at different X-Ray incidence angles: a) C1s, b) N1s, c) O1s. The theoretical spectra are reported in the upper part of each panel, while the experimental ones in the lower part. Vertical colored dashed bars:  $\Delta$ KS IPs. The calculated peaks have been shifted by -0.2 eV (a, c), and +0.45 eV (b) in order to facilitate the comparison with the experimental data.

From an inspection of Figure 16, we conclude that the theoretical results are able to reproduce the dichroic effects found in the experimental measurements. Indeed, in both C1s and N1s NEXAFS spectra, a progressive decrease of the intensity of spectral features is found by increasing the X-ray incidence angle; this is in line with the  $\pi$  nature of the final states responsible for the main spectral features.

The O1s spectra show the most intriguing trend: unlike the third peak, the first two peaks present the same dichroic behavior; indeed, by increasing the X-ray incidence angle, the intensity of the first two peaks increases, whereas the intensity of the third one decreases. The dichroic behavior of the computed absorption bands can be rationalized by considering the orientation of the molecule. Furthermore, the final states implicated in the most intense transitions of the first two peaks are mainly contributed by O  $2p_y$  AOs, while those relative to the third peak by O  $2p_z$  AOs; the direction of maximum amplitude of O  $2p_y$  and O  $2p_z$  AOs is perpendicular and parallel to the z axis, respectively. Bearing in mind that the intensity of a resonance is largest when the electric field vector lies along the direction of the final MO, and vanishes when it is perpendicular to it, one can conclude that the intensity of transitions giving rise to the first two peaks is largest at  $90^\circ$  and vanishes at  $0^\circ$ . Therefore, the similar dichroic behavior of the first two peaks is in line with the same nature of the final MOs involved in the most intense transitions (O  $2p_{xy}$  and  $\pi^*(\text{Ti}=\text{O})$ , respectively), while the opposite dichroic behavior of the third peak is due to the  $\sigma^*(\text{Ti}=\text{O})$  character of the final MO implicated in the main transition.

## Conclusions

A combined NEXAFS experimental and theoretical study on the electronic structure of TiOPc has been carried out. In particular, the aim was to determine the nature of the interaction between the molecule and the Ag(111) surface; to achieve this, NEXAFS spectra have been calculated at the C,N, and O *K-edge*, both for the randomly-oriented gas phase molecule and the fixed-in-space molecule, by means of DFT using the TP approximation. The calculated spectra have been compared to the experimental measurements, recorded at different X-ray incidence angles by the research group of professor G. Witte (University of Marburg, Germany) for a multilayer of TiOPc adsorbed on Ag(111). The nature of the adsorbate-substrate interaction has been analyzed by considering the differences between the spectral features relative to the gas phase and adsorbed molecule.

The analysis of NEXAFS spectra relative to the randomly-oriented gas phase molecule has revealed that both C1s and N1s spectral features in the below-edge energy region are due to transitions toward virtual MOs mainly contributed by out-of-plane C 2p and N 2p AOs.

The theoretical NEXAFS spectra calculated at different X-ray incidence angles reproduce the dichroic effects observed in the experimental measurements. In particular, a progressive decrease of the intensity of spectral features is found by increasing the X-ray incidence angle; in the C1s and N1s NEXAFS spectra, this trend reflects the  $\pi$  symmetry of the final states responsible for the main spectral features. As concerns the polarized O1s NEXAFS spectra, the same dichroic behavior is observed for the first two peaks, while the third one displays the opposite behavior. The similar dichroic behavior of the first two peaks is in line with the similar nature of the implicated final MOs (O 2p<sub>xy</sub> and  $\pi^*(\text{Ti}=\text{O})$ , respectively), whereas the opposite dichroic behavior of the third peak is due to the  $\sigma^*(\text{Ti}=\text{O})$  character of the final MO. To conclude, since a reasonable agreement has been found between theoretical calculations on the free molecule and experimental results for a multilayer of TiOPc adsorbed on Ag(111), the interaction of TiOPc with the Ag(111) surface is quite weak (*i.e.*, physisorption).

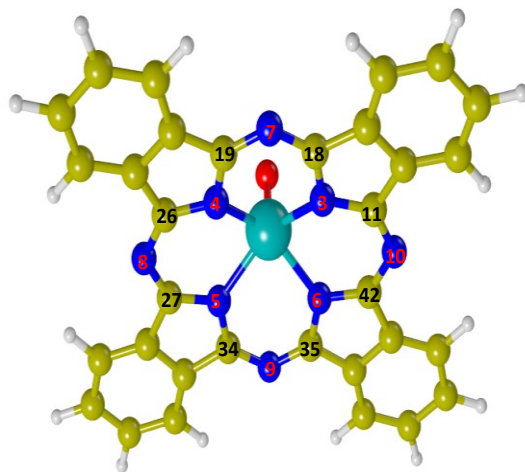
## Appendix

**Table 1** – Comparison between calculated and experimental geometrical parameters of TiOPc. The calculated parameters have been optimized at the DFT level by using the hybrid B3LYP xc potential. Bond lengths are expressed in Å, while angles in degrees.

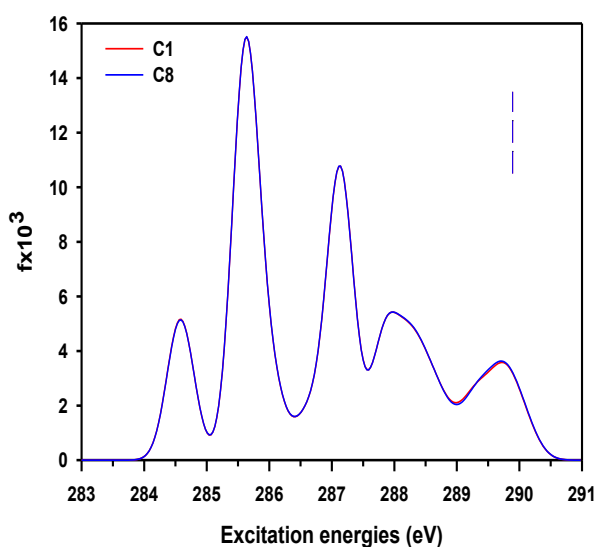
<b>Bond lengths</b>	Calculated	Experimental <sup>17</sup>
Ti=O	1.6295	1.6500
Ti-N5	2.0953	2.0530
Ti-N6	2.0955	2.0590
Ti-N3	2.0965	2.0840
Ti-N4	2.0947	2.0680
N5-C27	1.3764	1.3730
N5-C34	1.3760	1.3840
N6-C42	1.3762	1.3600
N6-C35	1.3767	1.3920

N3-C11	1.3764	1.3600
N3-C18	1.3764	1.3770
N4-C26	1.3762	1.3770
N4-C19	1.3762	1.3820
N8-C27	1.3243	1.3380
N8-C26	1.3241	1.3290
N9-C34	1.3244	1.3440
N9-C35	1.3239	1.3140
N10-C42	1.3246	1.3390
N10-C11	1.3244	1.3170
N7-C18	1.3239	1.3340
N7-C19	1.3241	1.3240
<b>Angles</b>	<b>Calculated</b>	<b>Experimental<sup>17</sup></b>
N5-Ti-N6	84.111	85.700
N4-Ti-N3	84.220	84.000
N5-Ti-N4	84.128	85.000
N6-Ti-N3	84.020	84.400
N4-Ti-N6	142.718	144.800
N3-Ti-N5	142.611	144.800
N4-Ti=O	108.650	108.900
N3-Ti=O	108.832	109.200
N6-Ti=O	108.632	106.300
N5-Ti=O	108.557	106.000

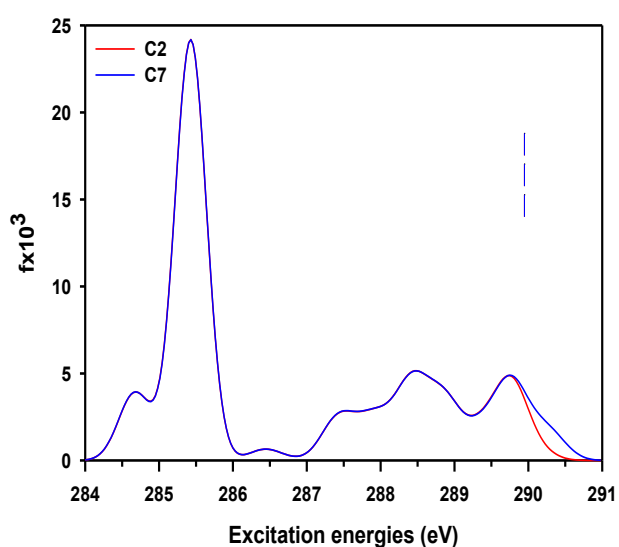




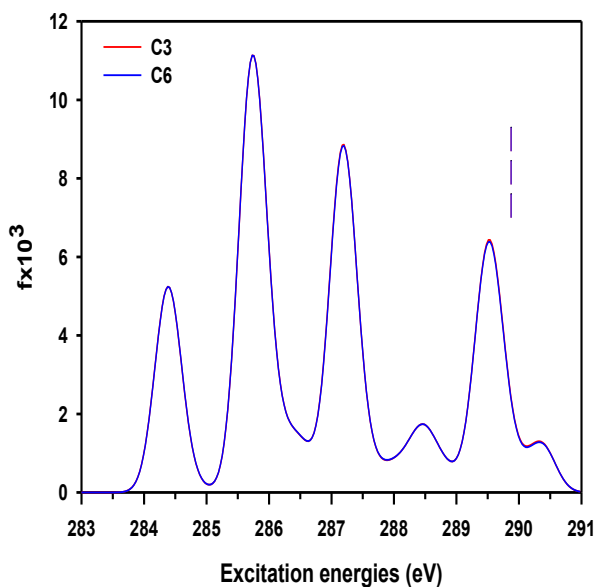
**Figure 1** – Chemical structure of TiOPc with the labeling of C and N atoms used in Table 1.



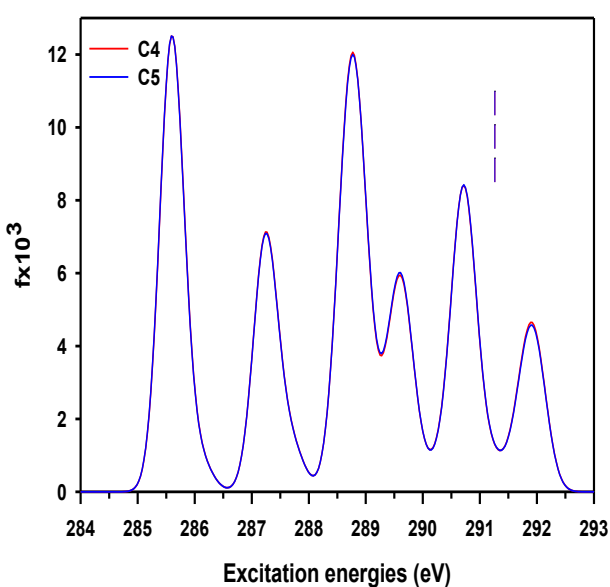
a)



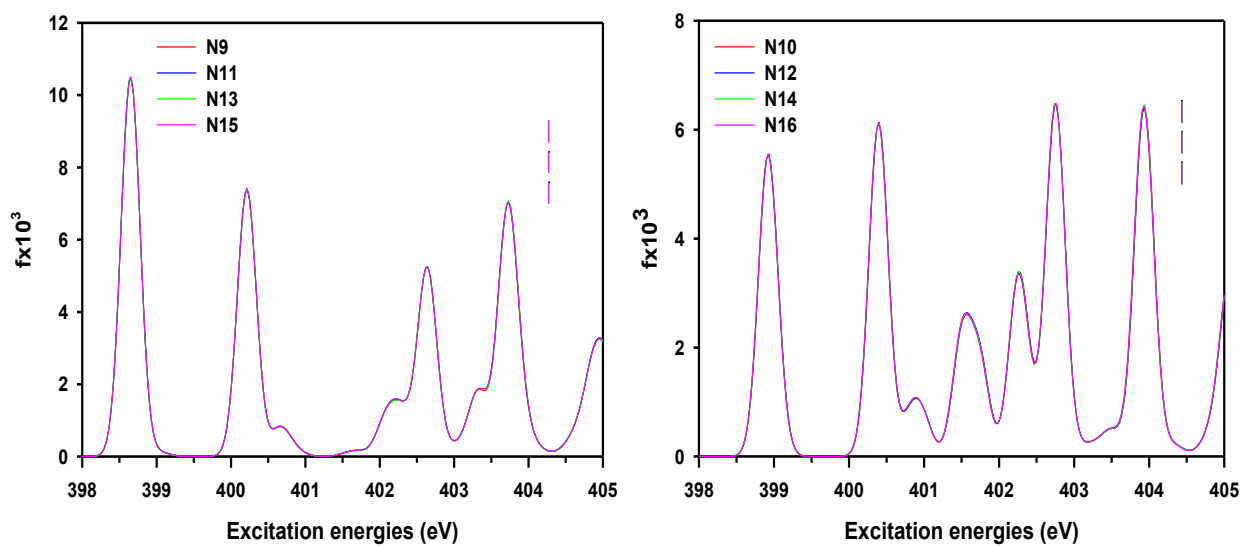
b)



c)

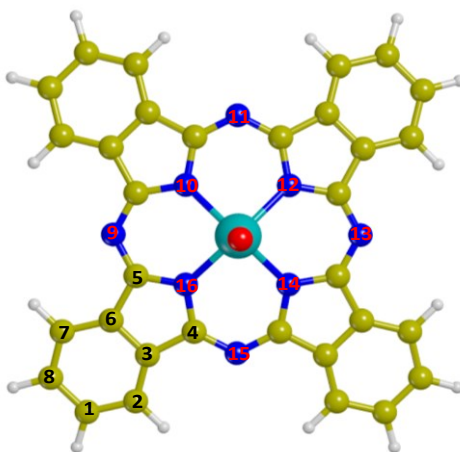


d)



e)

f)



g)

**Figure 2** – Panels a)-d): C1s NEXAFS spectral profiles of C1-C8, C2-C7, C3-C6, and C4-C5 couples, panels e), f): N1s NEXAFS spectral profiles of  $N_{aza}$  and  $N_{pyr}$  sites to check their equivalence. Vertical colored dashed bars:  $\Delta$ KS IPs. Panel g): chemical structure of TiOPc with the labeling of all C and N atoms considered in the tests.

**Table 2** – Peak assignments of the C1s NEXAFS spectrum of randomly-oriented gas phase TiOPc. Transition energies (eV), oscillator strengths ( $\text{fx}10^2$ ), and the composition of the involved MOs in terms of Mulliken population of AOs centered on each specific atom are reported. Only transitions with  $\text{fx}10^3 \geq 16.0$  are reported.  $C_h$  denotes the C atom carrying the core-hole.

Peaks	Core-hole site	E(eV)	$\text{fx}10^3$	Main atomic contribution to the final MO			Assignment
				Atom	AO	%	
<b>A</b>	C3	284.38	40.8	C	$2p_z$	61	$\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})$ (LUMO)
				N		27	
				$C_h$		4	
	C1	284.57	38.2	C	$2p_z$	62	
				N		27	
				$C_h$		4	
	C2	284.62	17.3	C	$2p_z$	62	
				N		28	
				$C_h$		2	
<b>B</b>	C2	285.43	193.0	C	$2p_z$	60	$\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})$
				$C_h$		23	
				N		6	
	C4	285.60	99.4	C	$2p_z$	51	$\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})$ (LUMO)
				N		30	
				$C_h$		9	
	C1	285.62	121.0	C	$2p_z$	69	$\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})$
				$C_h$		13	
				N		9	
				H		1	
				C		$2p_y$	
	C3	285.71	75.6	C	$2p_z$	71	$\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})$
				N		10	
				$C_h$		8	
	C3	285.90	18.2	C	$2p_z$	74	$\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})$
				N		18	
				$C_h$		2	
	C1	286.03	19.9	C	$2p_z$	77	$\pi^*(\text{C}=\text{C})+\pi^*(\text{C}-\text{N})$
				N		17	
				$C_h$		1	
	<b>C</b>	C1	287.13	85.1	C	$2p_z$	84
$C_h$					9		
N					8		
C3		287.18	67.4	C	$2p_z$	72	$\pi^*(\text{C}=\text{C})+3d \text{ Ti}$
				Ti	$3d_{xy}$	6	
				$C_h$		7	
				N	$2p_z$	4	

				O	2p <sub>y</sub>	2	
				Ti	3d <sub>yz</sub>	1	
				Ti	3d <sub>z<sup>2</sup></sub>	1	
	C4	287.24	52.8	C	2p <sub>z</sub>	62	$\pi^*(C=C)+\pi^*(C-N)+3d\ Ti$
				N		23	
				C <sub>h</sub>		5	
				Ti	3d <sub>xy</sub>	3	
<b>D</b>	C1	287.86	33.4	H	2s	33	$\sigma^*(C-H)$
				H	1s	30	
				C <sub>h</sub>	ns	38	
				Ti	1s	1	
				H	2p <sub>x</sub>	1	
				C	3p <sub>x</sub>	1	
				C	2p <sub>y</sub>	1	
	C2	287.92	16.1	C <sub>h</sub>	ns	53	$\sigma^*(C-H)$
				H	2s	25	
				H	1s	18	
				C	2p <sub>y</sub>	1	
	C1	288.21	19.1	C <sub>h</sub>	np <sub>y</sub>	19	$\sigma^*(C-H)+\sigma^*(Ti=O)$
				C <sub>h</sub>	np <sub>x</sub>	15	
				Ti	2s	14	
				Ti	3d <sub>z<sup>2</sup></sub>	12	
				H	1s	7	
				Ti	1s	5	
				Ti	2p <sub>z</sub>	4	
				O		4	
				H	2s	4	
				C <sub>h</sub>	ns	3	
				C <sub>h</sub>	3d <sub>xy</sub>	2	
	C4	288.61	16.5	Ti	3d <sub>yz</sub>	47	$\pi^*(Ti=O)+\pi^*(C=C)$
				C	2p <sub>z</sub>	20	
				O	2p <sub>y</sub>	14	
				Ti	3d <sub>xz</sub>	3	
				C <sub>h</sub>	2p <sub>z</sub>	2	
	C4	288.70	34.7	C	2p <sub>z</sub>	42	$\pi^*(C=C)+3d\ Ti$
Ti				3d <sub>xy</sub>	19		
Ti				3d <sub>x<sup>2</sup>-y<sup>2</sup></sub>	8		
Ti				3d <sub>yz</sub>	7		
C <sub>h</sub>				2p <sub>z</sub>	3		
N				2p <sub>x</sub>	2		
O				2p <sub>y</sub>	2		
N					1		
O				2p <sub>x</sub>	1		

				N	1s	1		
	C4	288.87	50.6	C	2p <sub>z</sub>	65	$\pi^*(C=C)+3d\ Ti$	
				Ti	3d <sub>xy</sub>	6		
				Ti	3d <sub>x<sup>2</sup>-y<sup>2</sup></sub>	5		
				Ti	3d <sub>yz</sub>	5		
				C <sub>h</sub>	2p <sub>z</sub>	4		
				Ti	3d <sub>xz</sub>	3		
				O	2p <sub>y</sub>	1		
				N	2p <sub>x</sub>	1		
<b>E</b>				C3	289.50	38.0		C
	N	25						
	C <sub>h</sub>	np <sub>z</sub>	12					
	C <sub>h</sub>	2p <sub>z</sub>	2					
	C4	289.61	40.5	C	2p <sub>z</sub>	70	$\pi^*(C=C)+\pi^*(C-N)+3d\ Ti$	
				N		17		
				Ti		3d <sub>xy</sub>		4
				Ti		3d <sub>x<sup>2</sup>-y<sup>2</sup></sub>		4
				C <sub>h</sub>		2p <sub>z</sub>		4
	C2	289.76	20.7	C	2p <sub>z</sub>	65	$\pi^*(C=C)+\pi^*(C-N)+ns,np\ C$	
				N		26		
				C <sub>h</sub>		np <sub>z</sub>		4
				C <sub>h</sub>		ns		4
				C <sub>h</sub>		2p <sub>z</sub>		1

**Table 3** – Peak assignments of the N1s NEXAFS spectrum of randomly-oriented gas phase TiOPc. Transition energies (eV), oscillator strengths (fx10<sup>2</sup>), and the composition of the involved MOs in terms of Mulliken population of AOs centered on each specific atom are reported. Only transitions with fx10<sup>3</sup> ≥ 5.00 are reported. N<sub>h</sub> denotes the N atom carrying the core-hole.

Peaks	Core-hole site	E(eV)	fx10 <sup>3</sup>	Main atomic contribution to the final MO			Assignment
				Atom	AO	%	
<b>A</b>	N <sub>aza</sub>	398.65	42.0	C	2p <sub>z</sub>	66	$\pi^*(C=C)+\pi^*(C-N)$ (LUMO)
				N <sub>h</sub>		8	
				N <sub>aza</sub>		7	
				N <sub>pyr</sub>		7	
	N <sub>pyr</sub>	398.91	21.1	C	2p <sub>z</sub>	68	$\pi^*(C=C)+\pi^*(C-N)$ (LUMO)
				N <sub>aza</sub>		13	
				N <sub>pyr</sub>		5	
				N <sub>h</sub>		4	
<b>B</b>	N <sub>aza</sub>	400.21	29.7	C	2p <sub>z</sub>	72	$\pi^*(C=C)+\pi^*(C-N)$
				N <sub>aza</sub>		17	

				N <sub>h</sub>		6	
	N <sub>pyr</sub>	400.39	24.5	C		69	$\pi^*(C=C)+\pi^*(C-N)$
				N <sub>pyr</sub>	2p <sub>z</sub>	9	
				Ti	3d <sub>xy</sub>	5	
				Ti	3d <sub>x<sup>2</sup>-y<sup>2</sup></sub>	3	
				N <sub>h</sub>	2p <sub>z</sub>	3	
				N <sub>pyr</sub>	2p <sub>y</sub>	1	
<b>C</b>				N <sub>pyr</sub>	401.51	8.87	
	O	2p <sub>x</sub>	13				
	C	2p <sub>z</sub>	12				
	Ti	3d <sub>yz</sub>	10				
	O	2p <sub>y</sub>	3				
	Ti	nd <sub>xz</sub>	2				
	Ti	3d <sub>z<sup>2</sup></sub>	1				
	N <sub>pyr</sub>	2p <sub>z</sub>	1				
	N <sub>pyr</sub>	401.74	6.58	Ti	3d <sub>xy</sub>	29	$\pi^*(C=C)+3d\ Ti+\sigma^*(C-N)$
				C	2p <sub>z</sub>	27	
				Ti	3d <sub>x<sup>2</sup>-y<sup>2</sup></sub>	16	
				N <sub>pyr</sub>	2p <sub>z</sub>	5	
				C	2p <sub>x</sub>	3	
				Ti	nd <sub>xy</sub>	2	
				N <sub>pyr</sub>	2p <sub>x</sub>	2	
				N <sub>pyr</sub>	2p <sub>y</sub>	2	
				N <sub>pyr</sub>	2s	1	
	N <sub>pyr</sub>	402.24	10.1	C		95	$\pi^*(C=C)+\pi^*(C-N)$
				N <sub>h</sub>	2p <sub>z</sub>	2	
				Ti	3d <sub>z<sup>2</sup></sub>	2	
	N <sub>aza</sub>	402.63	19.5	C		83	$\pi^*(C=C)+\pi^*(C-N)$
				N <sub>aza</sub>	2p <sub>z</sub>	10	
				N <sub>h</sub>		4	
				N <sub>h</sub>	np <sub>x</sub>	2	
				Ti	3d <sub>z<sup>2</sup></sub>	1	
	N <sub>pyr</sub>	402.75	25.4	C	2p <sub>z</sub>	71	$\pi^*(C=C)+\pi^*(C-N)+3d\ Ti$
				N <sub>pyr</sub>	2p <sub>z</sub>	13	
				Ti	3d <sub>xy</sub>	9	
Ti				3d <sub>x<sup>2</sup>-y<sup>2</sup></sub>	5		
N <sub>h</sub>				2p <sub>z</sub>	4		
<b>D</b>	N <sub>aza</sub>	403.32	6.31	N <sub>h</sub>	np <sub>z</sub>	83	Rydberg(N)+ $\pi^*(C=C)$
				Ti	3s	2	
				Ti	4s	2	
				C	np <sub>z</sub>	3	
				C	2p <sub>z</sub>	1	
	N <sub>aza</sub>	403.71	23.5	C	2p <sub>z</sub>	42	

				$N_h$	$3d_{xz}$	17	$\pi^*(C=C)+\pi^*(C-N)+3d,np N$
				$N_h$	$3d_{yz}$	1	
				$N_{aza}$	$2p_z$	14	
				$N_h$	$np_z$	5	
				$N_h$	$nd_{xz}$	3	
				$N_h$	$2p_z$	3	
				Ti	$3d_{z^2}$	1	
				$N_h$	$3d_{z^2}$	1	
				$N_h$	$3d_{x^2-y^2}$	1	
$N_{aza}$	403.83	5.55		$N_h$	$3d_{xz}$	41	Rydberg (N)+ $\pi^*(C=C)$
				C	$2p_z$	12	
				$N_h$	$nd_{xz}$	8	
				$N_h$	$3d_{yz}$	5	
				$N_h$	$np_z$	5	
				$N_{pyr}$	$2p_z$	3	
				Ti	3s	2	
				$N_h$	$2p_z$	2	
				$N_{aza}$	$2p_z$	2	
				$N_h$	$np_z$	1	
				$N_h$	$nd_{yz}$	1	
				Ti	4s	1	
$N_{pyr}$	403.94	22.4		C	$2p_z$	62	$\pi^*(C=C)+\pi^*(C-N)+3d N$
				$N_{aza}$	$2p_z$	15	
				$N_h$	$3d_{xz}$	10	
				$N_{pyr}$	$2p_z$	5	
				$N_h$	$nd_{xz}$	4	
				$N_h$	$p_x$	2	
				$N_h$	$np_z$	2	
				$N_h$	$2p_z$	3	
	$N_h$	$3d_{yz}$	2				

**Table 4** – Peak assignments of the O1s NEXAFS spectrum of randomly-oriented gas phase TiOPc. Transition energies (eV), oscillator strengths ( $\text{fx}10^3$ ), and the composition of the involved MOs in terms of Mulliken population of AOs centered on each specific atom are reported. Only transitions with  $\text{fx}10^3 \geq 1.00$  are reported.

Peaks	E(eV)	$\text{fx}10^3$	Main atomic contribution to the final MO			Assignment
			Atom	AO	%	
<i>A</i>	529.59	1.81	C	$2p_z$	65	$\pi^*(\text{C}=\text{C}) + \pi^*(\text{C}-\text{N})$ (LUMO+1)
			N		24	
			Ti	$3d_{xz}$	2	
<i>B</i>	530.99	20.4	Ti	$3d_{yz}$	60	$\pi^*(\text{Ti}=\text{O})$
			O	$2p_y$	11	
			Ti	$3d_{xz}$	5	
			N	$2p_z$	2	
			N	$2p_y$	2	
			Ti	$nd_{yz}$	2	
			Ti	$3d_{z^2}$	1	
	531.32	4.58	C	$2p_z$	48	$\sigma^*(\text{Ti}=\text{O}) + \pi^*(\text{C}=\text{C})$
			Ti	$3d_{z^2}$	21	
			O	$2p_z$	7	
			N		7	
	531.73	1.29	C	$2p_z$	80	$\pi^*(\text{Ti}=\text{O}) + \pi^*(\text{C}=\text{C})$
			Ti	$3d_{xz}$	5	
N			$2p_z$	5		
O			$2p_y$	4		
<i>C</i>	532.30	7.39	Ti	$3d_{z^2}$	37	$\sigma^*(\text{Ti}=\text{O}) + \pi^*(\text{C}=\text{C})$
			C	$2p_z$	37	
			O		11	
			N		2	
			O	$2s$	1	
			O	$ns$	1	



## Bibliography








- <sup>1</sup> L.-L. Chua, J. Zaumseil, J.-F. Chang, E.C.-W. Ou, P. K.-H.Ho, H. Sirringhaus, R.H. Friend, *Nature* **2005**, *434*, 194-199
- <sup>2</sup> M.A. Muccini, *Nat. Mater.* **2006**, *5*, 605-613
- <sup>3</sup> N. Karl, *Synth. Met.* **2003**, *133–134*, 649–657
- <sup>4</sup> T. Breuer, G. Witte, *Phys. Rev. B: Condens. Matter Mater. Phys.* **2011**, *83*, 155428
- <sup>5</sup> G. Schweicher, Y. Olivier, V. Lemaure, Y.H. Geerts, *Isr. J. Chem.* **2014**, *54*, 595–620
- <sup>6</sup> Y. Diao, K.M. Lenn, W.-Y. Lee, M.A. Blood-Forsythe, J. Xu, Y. Mao, Y. Kim, J.A. Reinspach, S. Park, A. Aspuru-Guzik, et al. *J. Am. Chem. Soc.* **2014**, *136*, 17046–17057
- <sup>7</sup> V. Oja, E.M. Suuberg, *J. Chem. Eng. Data* **1998**, *43*, 486–492
- <sup>8</sup> P. Rotter, B.A.J. Lechner, A. Morherr, D.M. Chisnall, D.J. Ward, A.P. Jardine, J. Ellis, W. Allison, B. Eckhardt, G. Witte, *Nat. Mater.* **2016**, *15*, 397–400
- <sup>9</sup> V. Mastryukov, C.-Y. Ruan, M. Fink, Z. Wang, R. Pachter, *J. Mol. Struct.* **2000**, *556*, 225-237
- <sup>10</sup> V. Oison, M. Koudia, M. Abel, L. Porte, *Phys. Rev. B* **2007**, *75*, 035428-1; 035428-6
- <sup>11</sup> P. Erk, H. Hengelsberg, in: K. M. Kadish, R. Guilard, K.M. Smith (Eds.), *The Porphyrin Handbook* Vol. 19, Amsterdam, **2003**, 105-150
- <sup>12</sup> K.Y. Law, *Chem. Rev. (Washington, D.C.)* **1993**, *93*, 449-486
- <sup>13</sup> D. Villemin, M. Hammadi, M. Hachemi, et al. *Molecules* **2001**, *6*, 831-844
- <sup>14</sup> A. Yamashita, S. Matsumoto, S. Sakata, T. Hayashi, and H. Kanbara, *J. Phys. Chem. B* **1998**, *102*, 5165-5167
- <sup>15</sup> S. Kera, M.B. Casu, K.R. Bauchspies, D. Batchelor, T. Schmidt, E. Umbach, *Surf. Sci.* **2006**, *600*, 1077-1084
- <sup>16</sup> H. Yamane, H. Honda, H. Fukagawa, M. Ohyama, Y. Hinuma, S. Kera, K. Okudaira, N. Ueno, *J. Electron Spectrosc. Relat. Phenom.* **2004**, *137–140*, 223–227
- <sup>17</sup> W. Hiller, J. Strähle, W. Kobel, M. Hanack, *Zeitschrift für Kristallographie* **1982**, *159*, 173-183
- <sup>18</sup> O. Okada, K. Oka, M. Iijima, *Jpn. J. Appl. Phys.* **1993**, *32*, 3556–3560
- <sup>19</sup> K. Oka, O. Okada, K. Nukada, *Jpn. J. Appl. Phys.* **1992**, *31*, 2181–2184
- <sup>20</sup> Z.D. Popovic et al., *J. Phys. Chem. B*, **1998**, *102* (4), 657–663
- <sup>21</sup> H. Yonehara, C. Pac, *Thin Solid Films* **1996**, *278*, 108-113
- <sup>22</sup> J.H. Park, L. Ravavar, I. Kwak, S.K. Fullerton-Shirey, P. Choudhury, A.C. Kummel, *J. Phys. Chem. C* **2017**, *121*, 6721–6728
- <sup>23</sup> Y. Alfredsson, H. Rensmo, A. Sandell, H. Siegbahn, *J. Electron Spectrosc. Relat. Phenom.* **2009**, *174*, 50–54

- 
- <sup>24</sup> J.H. Park, L. Ravavar, I. Kwak, S.K. Fullerton-Shirey, P.Choudhury, A.C. Kummel, *J. Phys. Chem. C* **2017**, *121*, 6721–6728
- <sup>25</sup> Y.Y. Wei, S.W. Robey, J.E. Reutt-Robey, *J. Phys. Chem. C* **2008**, *112*, 18537–18542
- <sup>26</sup> S. Colonna, G. Mattioli, P. Alippi, A.A. Bonapasta, A. Cricenti, F. Filippone, P. Gori, A.M. Paoletti, G. Pennesi, F. Ronci, G. Zanotti, *J. Phys. Chem. C* **2014**, *118*, 5255–5267
- <sup>27</sup> L. Fernández, S. Thussing, A. Mänz, G. Witte, A.X. Brion-Rios, P. Cabrera-Sanfeliix, D. Sanchez-Portal, P. Jakob, *J. Phys. Chem. C* **2017**, *121*, 1608–1617
- <sup>28</sup> I. Kröger, B. Stadtmüller, C. Kumpf, *New J. Phys.* **2016**, *18*, 113022
- <sup>29</sup> L. Fernandez, S. Thussing, A. Mänz, J. Sundermeyer, G. Witte, P. Jakob, *Phys. Chem. Chem. Phys.* **2017**, *19*, 2495–2502
- <sup>30</sup> M. Kothe, F. Widdascheck, G. Witte, *J. Phys. Chem. C* **2019**, *123*, 10, 6097–6106
- <sup>31</sup> R.G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules* **1989**, Oxford University Press, New York
- <sup>32</sup> A.D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652
- <sup>33</sup> C. Fonseca Guerra, J.G. Snijders, G. te Velde, E.J. Baerends, *Theor. Chem. Acc.* **1998**, *99*, 391-403
- <sup>34</sup> E.J. Baerends, D.E. Ellis, P. Roos, *Chem. Phys.* **1973**, *2*, 41-51
- <sup>35</sup> J.P. Perdew, *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *33*, 8822–8824
- <sup>36</sup> J.C. Slater, *Adv. Quantum Chem.* **1972**, *6*, 1-92.; L. Triguero, L.G.M. Pettersson, H. Ågren, *Phys. Rev. B* **1998**, *58*, 8097-8110
- <sup>37</sup> P. Norman, A. Dreuw, *Chem. Rev.* **2018**, *118*, 7208-7248
- <sup>38</sup> A. Mijiovilovich, L.G.M. Pettersson, S. Mangold, M. Janousch, J. Susini, M. Salome, F.M.F. deGroot, B.M. Weckhuysen, *J. Phys. Chem. A* **2009**, *113*, 2750-2756
- <sup>39</sup> R. De Francesco, M. Stener, G. Fronzoni, *J. Phys. Chem. A* **2012**, *116*, 2885–2894
- <sup>40</sup> G. Fronzoni, O. Baseggio, M. Stener, W. Hua, G. Tian, Y. Luo, B. Apicella, M. Alfè, M. de Simone, A. Kivimaki, M. Coreno, *J. Chem. Phys.* **2014**, *141*, 044313-1-044313-9

# Correlation effects in BIs core-excited states of boronic-acid derivatives: An experimental and computational study

Cite as: J. Chem. Phys. 151, 134306 (2019); <https://doi.org/10.1063/1.5120175>

Submitted: 16 July 2019 . Accepted: 12 September 2019 . Published Online: 02 October 2019

Daniele Toffoli , Aurora Ponzi , Elisa Bernes, Monica de Simone , Cesare Grazioli , Marcello Coreno , Matus Stredansky , Albano Cossaro, and Giovanna Fronzoni 



View Online



Export Citation



CrossMark

## Lock-in Amplifiers up to 600 MHz

starting at

\$6,210



Zurich  
Instruments

Watch the Video



# Correlation effects in B1s core-excited states of boronic-acid derivatives: An experimental and computational study

Cite as: *J. Chem. Phys.* **151**, 134306 (2019); doi: [10.1063/1.5120175](https://doi.org/10.1063/1.5120175)

Submitted: 16 July 2019 • Accepted: 12 September 2019 •

Published Online: 2 October 2019



View Online



Export Citation



CrossMark

Daniele Toffoli,<sup>1</sup> Aurora Ponzi,<sup>2</sup> Elisa Bernes,<sup>1</sup> Monica de Simone,<sup>3</sup> Cesare Grazioli,<sup>3</sup> Marcello Coreno,<sup>4</sup> Matus Stredansky,<sup>3,5</sup> Albano Cossaro,<sup>3</sup> and Giovanna Fronzoni<sup>1,a</sup>

## AFFILIATIONS

<sup>1</sup>Department of Chemical and Pharmaceutical Sciences, University of Trieste, via L. Giorgieri 1, 34127 Trieste, Italy

<sup>2</sup>Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia

<sup>3</sup>CNR-IOM Laboratorio TASC, Basovizza SS-14, km 163.5, 34149 Trieste, Italy

<sup>4</sup>CNR-I (LD2 Unit) Basovizza SS-14, Km 163,5 34149 Trieste, Italy

<sup>5</sup>Department of Physics, University of Trieste, via A. Valerio 2, 34127 Trieste, Italy

<sup>a</sup>E-mail: [fronzoni@units.it](mailto:fronzoni@units.it)

## ABSTRACT

We performed a theoretical investigation on the influence of electronic correlation effects on the B1s NEXAFS spectrum of boronic acid derivatives, namely, boric acid [B(OH)<sub>3</sub>], phenyl boronic acid (PBA), and 1,4-phenyl diboronic acid (PDBA), employing different computational schemes of increasing complexity, ranging from the purely one-electron scheme based on the transition potential method of density functional theory (DFT-TP), time-dependent DFT (TDDFT), and multiconfigurational self-consistent field (MCSCF). We also report experimental measurements of the B1s NEXAFS spectra of the aforementioned molecules together with the high-resolution C1s NEXAFS spectrum of PBA. We demonstrate that due to the shallow B1s core energy levels compared to C, O, and N, the inclusion of static correlation effects, which can be incorporated by using multireference approaches to excited states, assumes a decisive role in reconciling experiment and theory on B1s core-electron excitation energies and oscillator strengths to valence states. This claim is corroborated by the good agreement that we find between the DFT-TP calculated C1s NEXAFS spectrum and that experimentally measured for PBA and by the failure of both DFT-TP and TDDFT approaches with a selection of xc functionals kernels to properly describe the B1s NEXAFS spectrum of PBA and PDBA, at variance with the good agreement with the experiment that is found by employing the MCSCF wave function approach.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5120175>

## I. INTRODUCTION

Boronic acids are an important class of organoboron compounds with a wide range of applications in synthetic organic chemistry as well as in biology and medicine.<sup>1</sup> More recently, they have attracted special interest as building blocks in the synthesis of 2D boroxine-based systems.<sup>2-5</sup> In particular, the self-condensation of three boronic molecules leads to the formation of a boroxine group, a six membered ring of alternating oxygen and boron atoms. The use of phenyl precursors with two boronic terminations enables us to obtain extended covalent frameworks (COFs), whereas

molecules with a single boronic group can be adopted for the synthesis of macromolecules. In this context, phenyl-1,4-diboronic acid (PDBA) is widely used as a precursor of COFs,<sup>3</sup> while the simpler phenyl-boronic acid (PBA) with only one boronic group undergoes autocondensation to form triphenylboroxine (TPB) molecules.<sup>5</sup> The synthesis of TPB on the Au(111) surface was completely characterized in a recent work through a combined experimental and theoretical study of the boroxine-gold interface by means of STM (Scanning Tunneling Microscopy) and X-ray spectroscopies.<sup>5</sup> The inclusion in the computational scheme of the relaxation effects following the core hole formation is essential for a correct reproduction

of the spectral features and was achieved by means of the transition potential (TP) approximation.<sup>6,7</sup> This well-established technique generally works well for describing K-shell excitations of the first row atoms of the periodic table in systems of varying dimensionality and is able to correctly describe both the energy separation and the intensity distribution among the spectral features. However, the application of the DFT-TP scheme to the description of the B1s core excitations of the aforementioned boronic systems [free TPB and TPB@Au(111) model cluster] showed an unexpected discrepancy between experiments and theoretical results.<sup>5</sup> In particular, the calculations overestimate by about 0.6 eV the energy separation between the two main peaks of the B1s spectrum and do not correctly reproduce the intensity distribution between them, while the O1s experimental NEXAFS spectrum is correctly reproduced by the theory. Calculations performed by employing different flavors of exchange-correlation (xc) potentials at both the DFT and TDDFT did not improve the agreement with the experimental data. We therefore ascribed the observed discrepancies to the limited amount (or the lack) of final state correlation effects (in particular, static correlation effects) included in the computational approaches.<sup>5</sup>

We consider it important to analyze in depth the role played by electron correlation effects in the B1s NEXAFS spectrum simulation of a class of boronic acid derivatives, in view of their significant importance in the realization of COFs and of their accurate electronic characterization to understand the intermolecular interactions that give rise to the networks on metal surfaces. This goal can be achieved by considering the calculation of the B1s core spectra on passing from a purely one-electron picture, such as the DFT-TP scheme, to explicitly correlated wave function approaches. The inclusion of high order correlation effects limits the applicability of this analysis to only rather small systems because of the rapid increase in the computational effort with the system dimension.<sup>8</sup> The understanding of the effects on model molecules would make it possible to evaluate possible discrepancies between calculated and experimental spectra of extended systems, specifically large boronic networks, for which calculations beyond the one-particle scheme become quickly unfeasible.

In this work, we systematically investigate the influence of electron-correlation effects on the B1s NEXAFS spectrum of a series of isolated boronic molecules of increasing complexity, namely, B(OH)<sub>3</sub>, PBA, and PDDBA, the latter being a commonly used precursor of boroxine-based networks. Different computational schemes of increasing complexity are employed in order to reproduce the high resolution NEXAFS measurements performed at the B1s edge of the gas phase samples which provide a direct and strict test of the quality of the calculated excitation energies and oscillator strengths. To our knowledge, the electronic structure characterization of these boronic-acid derivatives through a combined experimental and theoretical analysis of their B1s NEXAFS spectra has not been reported yet.

## II. EXPERIMENTAL DETAILS

The measurements were performed at the Gas Phase Photoemission beam line of the Elettra Synchrotron in Trieste (Italy).<sup>9</sup> The samples were vaporized in vacuum at temperatures rather below the respective melting points, and exactly 78 °C for B(OH)<sub>3</sub>, 36 °C for PBA, and below 120 °C for PDDBA, using an

inductively heated oven. X-ray Absorption Spectra (XAS) at boron and carbon K-edges were acquired by measuring the Total Ion Yield (TIY) with an electron multiplier placed in the experimental chamber in front of the ionization center. The exciting photon flux was measured simultaneously using a calibrated Si photodiode (AxVU100 IRD<sup>®</sup>) for normalization of the spectra. The energy scale of the spectra was calibrated by taking simultaneous spectra of the samples and of the reference gas: the second order signal of N<sub>2</sub>, N1s → π, ν = 1: 401.1 eV/2 = 200.55 eV for the boron K-edge<sup>10</sup> and the carbon core absorption spectrum of CO<sub>2</sub> for C K-edge. The photon resolution of spectra was better than 25 meV for boron K-edge and around 150 meV for carbon one. X-ray Photoemission Spectra (XPS) data were acquired using a Scienta SES-200 electron analyser<sup>11</sup> mounted at the magic angle with respect to the electric vector of the linearly polarized incident light. B1s XPS spectra have been recorded at 262 eV with an overall resolution better than 130 meV and have been calibrated with respect to S2p lines (180.2 and 181.5 eV) of SF<sub>6</sub>.<sup>12</sup> C1s XPS spectra have been measured at 382 eV, with an overall resolution of about 130 meV, and calibrated with respect to C1s of CO<sub>2</sub> at 297.70 eV.<sup>13</sup>

## III. THEORETICAL METHODS AND COMPUTATIONAL DETAILS

The equilibrium geometry of the three molecules was computed at the local density approximation (LDA) DFT level by employing the Vosko-Wilk Nusair<sup>14</sup> xc potential and a basis set of Slater-type orbitals (STO) of triple-zeta quality (TZP) for all atoms. Calculations were done by using the ADF quantum chemistry program.<sup>15</sup> The equilibrium geometries of the three molecules are included in Table S1 of the [supplementary material](#). In agreement with a recent work by Rao *et al.*,<sup>16</sup> the lowest energy conformer of PBA is the planar endo-exo. For PDDBA, we find an equilibrium geometry similar to that of Pawlak *et al.*<sup>17</sup> Optimized geometries yield structural parameters (bond distances and bond angles) in very close agreement with those derived from the experimentally determined crystalline structure for all the molecules considered, as shown in Tables S2–S4 of the [supplementary material](#). The B1s NEXAFS calculations for the three molecules were first computed at the DFT level within the Generalized Gradient Approximation (GGA) with the PW86xPerdew (PW86) xc functional<sup>18</sup> and the TP scheme (DFT-TP).<sup>6,7</sup> In the DFT-TP computational technique, half an electron is removed from the 1s orbital of the B atom, relaxing all the orbitals until self-consistency is obtained. Relaxation effects upon formation of the core hole are usually adequately included in this approach.<sup>19</sup> To accurately describe excitations to diffuse Rydberg states close to the ionization threshold, an even tempered quadruple-zeta with a three polarization and three diffuse function (QZ3P-3DIF) basis set was employed for the core-excited B atom, while a TZP basis set was used for the remaining atoms. Taking PBA molecule as a test case, the performance of other DFT functionals, namely, the hybrid B3LYP<sup>20</sup> and the meta-hybrid M06 and meta-GGA M06-L,<sup>21</sup> was first tested. The B1s core excitations of PBA were also calculated employing the TDDFT approach with the exchange correlation kernel approximated according to the Adiabatic Local-Density Approximation (ALDA).<sup>22</sup> At the TDDFT level, the coupling between the single excited configurations is formally included. The core-excited states can be computed efficiently by restricting the

single excitation space to include only excitations from the subset of core orbitals.<sup>23</sup> The TDDFT core excitation energies are generally underestimated due to the approximate exchange within the exchange-correlation functionals; to correct this error, exchange correlation functionals with varying fractions of Hartree-Fock (HF) exchange were considered. In particular, the following xc potentials with the ground state electron configuration were employed: the standard LB94 functional,<sup>24</sup> characterized by the correct asymptotic behavior, the hybrid B3LYP and PBE0,<sup>25</sup> and the meta-hybrid xc functional M06. The TDDFT simulations of the B1s spectra of PDDBA and B(OH)<sub>3</sub> were also performed, and the effects of the different xc functionals were analyzed.

*Ab initio* correlated wave function approaches to the calculation of the B1s NEXAFS spectra of PBA and PDDBA were done using the Multi-Configuration Self-Consistent-Field (MCSCF) method<sup>26,27</sup> as implemented in the MOLPRO quantum chemistry package.<sup>28</sup>

MCSCF calculations at the DFT LDA equilibrium geometry employ Dunning's correlation-consistent basis sets: an aug-cc-pVTZ basis for O and C atoms, a cc-pVTZ basis set for H atoms, and a doubly augmented basis set (d-aug-cc-pVTZ) for the B atom. We follow a computational procedure similar to that of Ref. 29, with some important modifications, due to the much larger size of the systems under study. The ground and core-excited states are calculated separately in two different MO basis sets while transition moments are obtained by means of a biorthogonalization procedure. Since the localization of the core-hole does not lower the symmetry of the molecules, all calculations employ the full C<sub>s</sub> symmetry. Considering first the PBA molecule, the active space employed for the MCSCF calculations comprises 15 active orbitals and 12 electrons. The set of active orbitals comprises 7 orbitals of a'' symmetry (2 occupied and 5 virtual orbitals) and 8 orbitals of a' symmetry (4 occupied orbitals and 4 virtual orbitals). To determine the ground-state wave function, an HF run is followed by a MCSCF calculation on this active space. The B1s orbital was included in the active space, but kept frozen and doubly occupied. This is followed by a configuration-interaction (CI) calculation on the reference-space only, with the same active space. This last step is needed for the calculation of the transition moments with the core-excited states. For the calculation of the core-excited states, a second MCSCF run was performed with the same active orbital space, with the B1s orbital kept frozen and its occupation restricted to one electron. The MCSCF calculations are carried out in the state-averaged mode (10 states) with equal weight for all states. The same procedure is also followed for the calculation of the core-excited states of PDDBA: the active space used in the MCSCF calculations comprises 14 active orbitals (6 orbitals of a' symmetry and 8 orbitals of a'' symmetry) and 12 electrons. The choice of the basis set and active spaces was done after several attempts and tests on both molecules and gave the most accurate results in terms of energy positions and intensity distribution of the first two absorption bands assigned to excited states of valence character.

For the C K-edge NEXAFS spectrum calculation of PBA, a separate computation of the excitation spectrum of each nonequivalent C site has been performed, employing the DFT-TP method with the PW86 functional. The total spectrum has been obtained by summing up the different contributions. The even tempered QZ3P-3DIF basis set has been used for the core-excited C atom, while a TZP basis

set has been employed for the remaining atoms; core orbitals of the latter atoms have been treated by the Frozen Core (FC) technique ([TZP.1s] basis set in the ADF database).

In the figures of each DFT-TP and TDDFT calculated spectrum, the B1s/C1s ionization thresholds are shown and used to separate the below- and the above-edge regions. Only transitions that occur below the ionization threshold can be accurately described by the employed computational protocol, while above it only qualitative information can be extracted, since the electronic continuum wave function cannot be properly described with the standard basis sets of quantum chemistry programs.

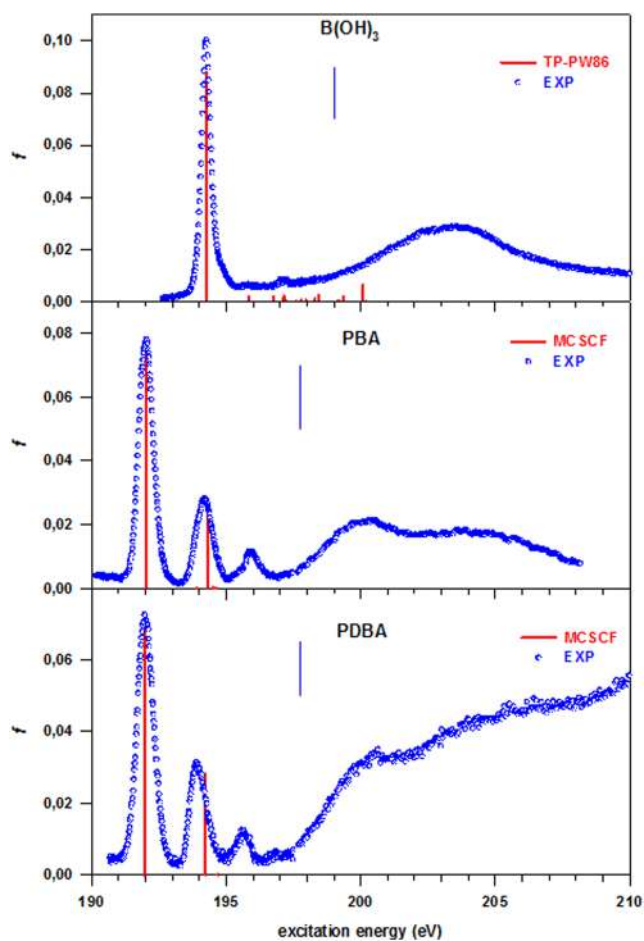
All the calculated stick spectra have been broadened by using a Gaussian lineshape  $\Phi(\omega; \Delta\epsilon_{1s} \rightarrow a, \gamma)$  with Full-Width-at-Half-Maximum (FWHM)  $\gamma = 0.5$  eV for B1s and  $\gamma = 0.3$  eV for C1s.

#### IV. RESULTS AND DISCUSSION

The B1s experimental NEXAFS spectra of B(OH)<sub>3</sub>, PBA, and PDDBA are shown in Fig. 1 (see molecular pictures in Tables S2–S4 of the [supplementary material](#); the theoretical results, in terms of B1s core excitation energies and oscillator strengths, which provide the best match with the experimental data, are also reported, in particular, the DFT-TP results for B(OH)<sub>3</sub> and MCSCF results for PBA and PDDBA. The B(OH)<sub>3</sub> spectrum shows only one strong peak below the ionization threshold while the spectra of PBA and PDDBA show a very similar intensity distribution characterized by a main peak followed by two features of decreasing intensity. The position of the first peak is significantly shifted to lower energies (by about 2.5 eV) in going from B(OH)<sub>3</sub> to PBA, while it does not change further from PBA to PDDBA. The experimental energies were extracted by fitting the data with a Gaussian curve for each main visible structure in the spectrum. The B(OH)<sub>3</sub> spectrum shows only one strong peak below the ionization threshold at 194.27 eV. This peak is assigned to the transition to the  $1\pi^*$  orbital provided by the calculations. Since the main peak shows a slight asymmetry attributed to vibrations, two peaks with much lower intensity (at 193.85 and 194.85 eV) were provided in the fit to account also for the vibrational progression. At higher photon energies, two weak features are also detected (at 195.74 and 197.02 eV) having Rydberg character.

The spectra of PBA and PDDBA are very similar to each other but differ from B(OH)<sub>3</sub>: they are broader than B(OH)<sub>3</sub> and consist in a main peak followed by two peaks of decreasing intensity. Given the broader width, it was not possible to appreciate by fitting if there are vibrational structures also for PBA and PDDBA.

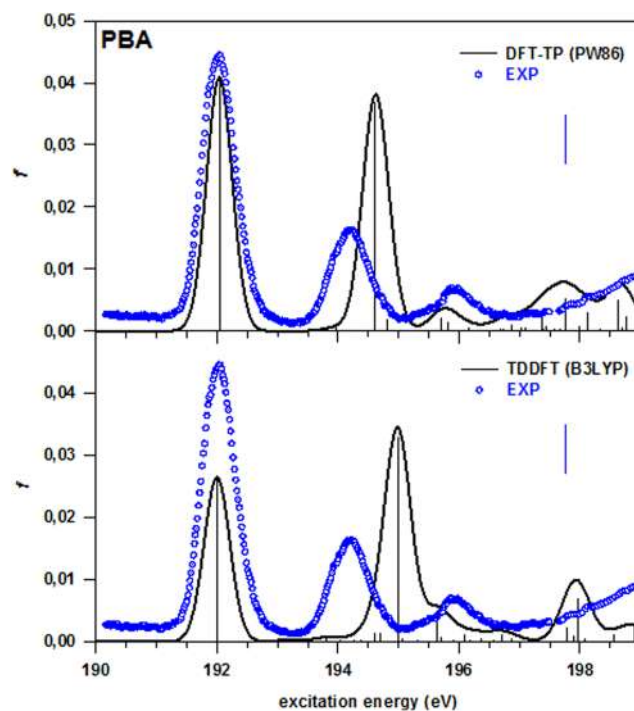
In order to have a key for the interpretation of the core spectra in terms of a single particle model, it can be useful to consider the theoretical results obtained at the DFT-TP level with the PW86 functional for the three molecules [reported in Table S5 of the [supplementary material](#) together with the experimental energies and in Figs. 1, 2, and 4 for B(OH)<sub>3</sub>, PBA and PDDBA, respectively]. Also plots of the MOs of  $\pi^*$  nature involved in the most intense transitions giving rise to the first two peaks of PBA and PDDBA spectra are reported in Fig. S1 of the [supplementary material](#). The low-lying intense peak of the three spectra is assigned to the B1s  $\rightarrow$  LUMO transition. The LUMO orbital is an antibonding  $\pi^*$  orbital contributed by the 2p perpendicular atomic components of B and, to a lesser extent, O atoms in B(OH)<sub>3</sub>, which also delocalizes on the phenyl ring bonded to the B atom in PBA and PDDBA molecules.



**FIG. 1.** Experimental B 1s NEXAFS spectra of  $B(OH)_3$  (upper panel), PBA (central panel), and PDBA (lower panel). Vertical red lines: calculated excitation energies and oscillator strengths at the DFT-TP level for  $B(OH)_3$  and at the MCSCF level for PBA and PDBA. A rigid shift to lower excitation energies of  $-0.47$  eV,  $-2.48$  eV, and  $-2.26$  eV was applied to the calculated excitation energies of  $B(OH)_3$ , PBA, and PDBA, respectively, to align the first excitation energy to the first experimental peak. Vertical blue lines: IP experimental values [199.0 eV, 197.75 eV, and 197.75 eV for  $B(OH)_3$ , PBA, and PDBA, respectively].

The stabilizing effect due to the aromaticity and the additional relaxation induced by the phenyl ring accounts for the lowering of the first excitation energy on passing from  $B(OH)_3$  to the two boronic acid derivatives. The higher energy low intensity lines of the  $B(OH)_3$  spectrum below edge (in the energy range 195–199 eV) correspond to transitions into MOs of diffuse nature and correctly reproduce the spectral shape at the higher energy side of the main peak (see the upper panel of Fig. 1). In order to facilitate the comparison with the experiment, Fig. S6 of the [supplementary material](#) shows the DFT-TP B1s stick spectrum of figure broadened by using a proper Gaussian lineshape.

The calculated energy and intensity of the second peak, visible in the experimental B1s NEXAFS spectra of both PBA and PDBA, are instead overestimated by the DFT-TP calculations with



**FIG. 2.** Comparison between calculated and experimental NEXAFS B1s spectra of PBA. Upper panel: DFT-TP results with the PW86 xc functional. Lower panel: TDDFT results with the B3LYP xc functional. The calculated spectra were rigidly shifted on the experimental energy scale (DFT-TP:  $+0.04$  eV; TDDFT:  $+10.16$  eV). Vertical blue lines: IP experimental values.

the PW86 functional (see Figs. 2 and 4): it has again a  $\pi^*$  nature in both spectra, very similar to the first more intense peak, associated with the 2p AOs perpendicular to the molecular plane of carbons in the phenyl ring and of the B atom bonded to the ring. The third less intense feature (around 196 eV) in the PBA and PDBA spectra is instead assigned to empty molecular orbitals of diffuse character. The DFT-TP failure in describing the B1s core excitations of PBA and PDBA molecules was investigated in depth by employing different theoretical approaches, both in DFT and TDDFT frameworks and in the multireference MCSCF level of theory that includes higher order (mainly static) electron correlation effects. To this purpose, we focus on the PBA molecule as a test case. The most accurate results for PBA, obtained employing the PW86 functional at the DFT-TP level and the B3LYP functional at the TDDFT level, are reported in Fig. 2 together with the experimental spectrum for comparison.

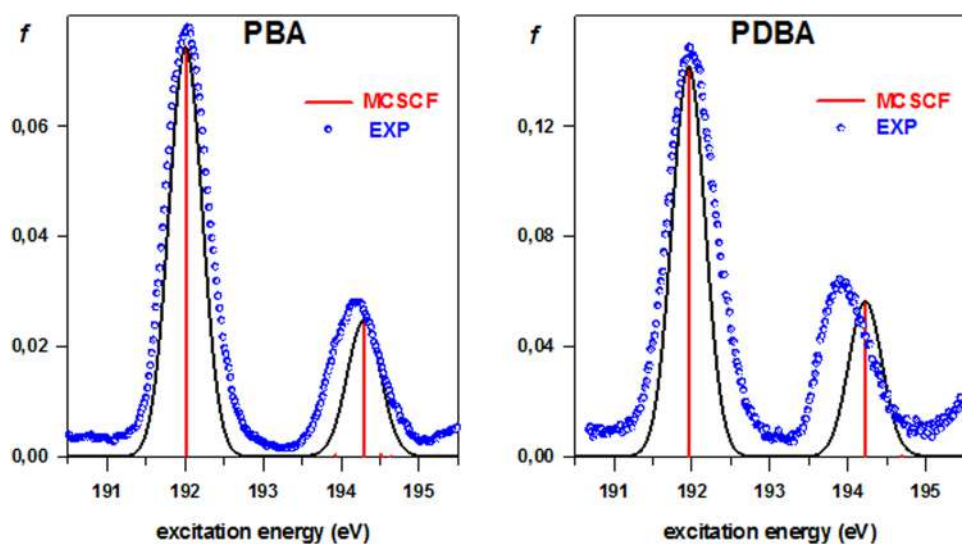
The DFT-TP spectrum shows two main deficiencies with respect to the experiment: the overestimation (of about 0.4 eV) of the energy separation between the first two peaks as well as an incorrect intensity distribution, with very similar oscillator strengths calculated for the two peaks. The effect of the potential choice on the DFT-TP results was assessed by employing other exchange-correlation potentials (hybrid B3LYP, and meta-hybrid M06 and M06-2x). The discrepancies with the experiment persist however in all the simulations (see Fig. S2 of the [supplementary material](#)), suggesting that

such poor performance is due to the neglect of electron-correlation effects that are beyond DFT.

It is important to observe that all calculations confirm the same nature for the calculated features, independently from the choice of the xc potential. To go beyond the one-electron approximation and include explicitly many-body effects in the computational approach, we first considered the TDDFT method, which formally includes the coupling between the single excited (1h-1p) configurations. The lower panel of Fig. 2 compares the TDDFT (B3LYP) results with the experiment. Focusing on the first two peaks, a further deterioration of the description is apparent compared to the DFT-TP results, with an additional increase in the energy separation (0.82 eV) between the two low-energy absorption bands, as well as of the intensity of the second peak that now exceeds that of the first one. As concerns the nature of the excitations, the lower lying final state is described by a single 1h-1p configuration corresponding to a pure B1s to LUMO transition in line with the DFT-TP description, while the second peak is contributed by two excitations (at 195.02 eV and a less intense at 195.66 eV) deriving from the mixing of two configurations. These involve with different weight two B1s transitions toward the same final orbitals, in particular, a higher energy  $\pi^*$  valence MO, delocalized over the phenyl ring and the boronic group, and a quite diffuse MO still of  $\pi^*$  symmetry. No significant variations of the energy splitting and the intensities distribution among the first two peaks were obtained employing other functionals (LB94, PBE0, and M06), as shown in Fig. S3 of the [supplementary material](#). It is interesting to underline that the TDDFT calculations performed on the B1s spectrum of B(OH)<sub>3</sub> confirm instead the spectral description obtained at the DFT-TP level. The B(OH)<sub>3</sub> TDDFT results, obtained with different xc functionals, are reported in the Fig. S7 of the [supplementary material](#), which shows the good reproduction and the substantial convergence toward the experimental profile of all the DFT based calculations. The analysis of all these results suggests that the configuration mixing described by the TDDFT with the ALDA approximation is not adequate to properly describe the B1s spectrum of PBA. It is by now well known that the ALDA

approximation for the xc functional within linear response TDDFT fails in describing double excitations;<sup>30-32</sup> therefore, it appeared necessary to employ *ab initio* techniques able to include electronic excitations higher than 1h1p in the configuration space. In this respect, a well-established approach to treat many-body effects in excited states is the MCSCF, which allows us to construct a reasonable number of configurations by considering all the possible excitations among the orbitals of a selected active space. The active space employed for the MCSCF calculations includes some virtual MOs with diffuse character which are however not sufficient to describe accurately the higher energy states generated by the MCSCF procedure starting from this active space. Therefore, it proves appropriate for an accurate description of the lowest lying core excited states mostly of valence character, while higher energy excitations of diffuse nature are not correctly taken into account. The use of larger active spaces is however intractable given the size of the systems at hand. However, the low-energy region of the NEXAFS B1s spectrum is anyway the more interesting one, where the breakdown of DFT theories is more evident. Accordingly, Fig. 3 reports only the lowest lying calculated core electron excitations which account for the first two bands observed in the NEXAFS experimental spectrum of PBA (the left panel of Fig. 3).

A significantly better description of the first two bands of the B1s spectrum is now obtained, with an almost quantitative agreement with the experiment both as concerns the energy separation between the two peaks and their relative intensity. The actual reason for such an improvement can be identified inspecting the nature of the MCSCF wave functions. The lowest lying transition (calculated at 194.51 eV), associated with the first peak in the experiment at 192.03 eV (see Fig. 3 and Table S5), is well described by a single 1h-1p configuration (its coefficient is 0.91) corresponding to the B1s  $\rightarrow$  6a'' (LUMO) transition. Its calculated intensity is the largest in agreement with the experimental intensity trend. The next calculated excitation (at 196.80 eV), describing the second experimental peak at 194.20 eV (see Fig. 3 and Table S5), is actually described by two strongly mixed configurations: the 1h1p

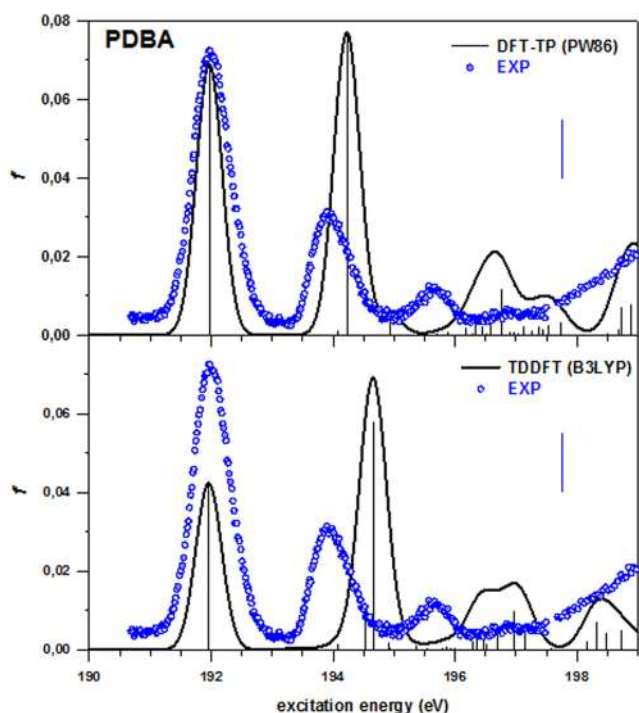


**FIG. 3.** Comparison between theoretical MCSCF and experimental spectra of the lower lying B1s core excitations of PBA (left panel) and PDBA (right panel) of valence nature. The calculated spectra were rigidly shifted on the experimental energy scale ( $-2.48$  eV for PBA;  $-2.26$  eV for PDBA).



$B1s \rightarrow 8a''$  (LUMO+2) excitation (with coefficient  $-0.521$ ) and the  $2h2p$  excitation [ $B1s^{-1}(\text{HOMO}-1)^{-1}(6a'')^2$ ] (with coefficient  $0.588$ ) which involves the simultaneous excitation of the  $B1s$  core electron and of a second electron from the  $\text{HOMO}-1$  occupied MO into the LUMO orbital. The calculated oscillator strength for this transition is lower than that predicted by DFT-TP and TDDFT and correctly reproduces the experimental intensity distribution between the two peaks. The presence of a doubly excited configuration in the description of this excited state has therefore a profound effect on the description of the second peak in the  $B1s$  NEXAFS spectrum of PBA and accounts for the failure of both the single particle (DFT) and the  $1h1p$  configuration-interaction (TDDFT) approaches. It is worth noting that the final orbitals ( $6a''$  and  $8a''$ ), involved in the two excited configurations that describe the transition calculated at  $196.80$  eV, both have a  $\pi^*$  character. The diffuse MOs included in the active space do not contribute to these excited configurations, further confirming the  $\pi^*$  nature of the second experimental peak, in line with the DFT-TP attribution. This observation is important in the perspective of a  $B1s$  NEXAFS spectral interpretation based on the DFT-TP approach of large systems for which a MCSCF approach like the one used here quickly becomes unfeasible.

The DFT-TP (with the PW86 functional) and TDDFT (with the B3LYP functional) approaches also fail when applied to the  $B1s$  spectrum simulation of the PDBA molecule, as it is clearly visible in Fig. 4 where the results are compared with the experimental spectrum.



**FIG. 4.** Comparison between calculated and experimental NEXAFS  $B1s$  spectra of PDBA. Upper panel: DFT-TP results with the PW86 xc potential. Lower panel: TDDFT results with the B3LYP xc potential. The calculated spectra were rigidly shifted on the experimental energy scale (DFT-TP:  $0.19$  eV; TDDFT:  $+10.41$  eV). Vertical blue lines: IP experimental values.

Apart from the overestimation of the energy separation between the two lower lying peaks, a wrong intensity distribution between them is also provided by both the computational schemes.

The influence of different DFT exchange-correlation potentials was again tested. Bearing in mind what the study on PBA pointed out, we restricted the performance analysis at the DFT-TP level to only two xc functionals (PW86 and B3LYP), while all the xc functionals employed for the PBA molecule were considered in the TDDFT calculations. The results are collected in Figs. S4 and S5 of the [supplementary material](#).

The inspection of the figures indicates that the disagreement between theory and experiment cannot be removed by simply varying the xc energy functional, similar to what was already found for PBA. Concerning the interpretation of the main structures of the PDBA  $B1s$  spectrum, it is convenient to rely on the DFT-TP results of Fig. 4 (upper panel), pointing out that the interpretation does not change irrespective of the xc potential employed.

Both the first two peaks have a  $\pi^*$  nature and the relative final orbitals (LUMO and LUMO+3) involve the phenyl ring as well as the  $2p$  AOs of the B atoms and, to a minor extent, of the O atoms perpendicular to the molecular plane. The less intense feature at higher energy is ascribed to empty orbitals of diffuse nature.

The TDDFT-B3LYP results (the lower panel of Fig. 4) are in even worse agreement with the experiment, in line with the trend found for PBA. The lower lying absorption peak is described by a single  $1h-1p$  configuration corresponding to the  $B1s \rightarrow \text{LUMO}$  excitation while the second one is contributed by two lines (calculated at  $184.18$  eV and  $184.33$  eV) with a very different intensity. Each one is characterized by a combination with different weights of the same two  $1h1p$  excited configurations involving the excitations of the  $B1s$  core electron into the second  $\pi^*$  valence MO and into a higher energy diffuse MO. The configuration mixing restricted to the  $1h1p$  configuration space proves to be not adequate to correctly describe the  $B1s$  core excitations also in the PDBA molecule. A substantially improved agreement with the experiment is obtained at the MCSCF level, as illustrated in the right panel of Fig. 3. The energy separation between the two calculated bands is now well described although a slight overestimate is still present, while their intensity ratio is very satisfactory. At the MCSCF level, the core hole final state (calculated at  $194.23$  eV) and associated with the lowest lying band in the experiment at  $192.00$  eV (see Fig. 3 and Table S5) is well described by a single configuration (the relative coefficient is  $0.91$ ) corresponding to the  $B1s \rightarrow 8a''$  (LUMO) excitation. The second final state (calculated at  $196.49$  eV) attributed to the second experimental peak at  $193.97$  eV (see Fig. 3 and Table S5) instead has a multiconfiguration nature being essentially described by two dominant configurations: the single excitation  $B1s \rightarrow 10a''$  (LUMO+2) and the doubly excited configuration  $B1s^{-1}(\text{HOMO}-1)^{-1} \rightarrow 8a''$  whose CI coefficients are  $-0.59$  and  $0.62$ , respectively. Both virtual orbitals are of  $\pi^*$  valence nature. In summary, the analysis of the MCSCF results indicates the importance of doubly excited configurations in the description of  $B1s$  core-excited states in PBA and PDBA boronic acid derivatives.

In order to ensure that the aforementioned many-body effects are a specific feature of the  $B1s$  core excited states of these systems, and to a much lesser extent, the K-shell excitation of the heavier C and O elements, we decided to calculate the  $C1s$  NEXAFS spectrum of PBA, for which we also report here the high resolution XAS

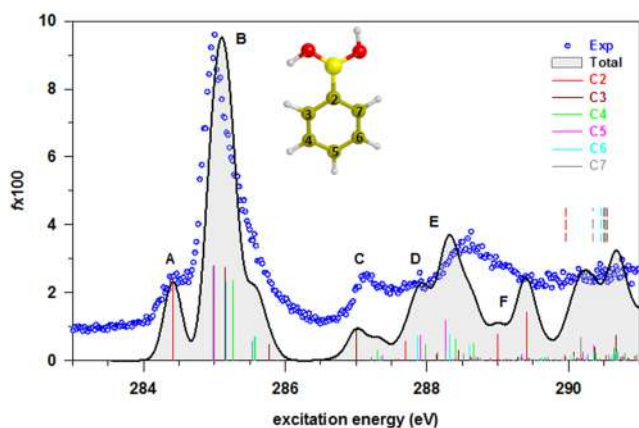
spectrum at the C K-edge, essential for a meaningful comparison with the theoretical data. The calculations of the C1s NEXAFS spectrum of PBA were performed at the DFT-TP level, bearing in mind that this approach is very effective to describe this edge in a variety of molecular systems, as cited in the introduction, and gave results in very good agreement with experiments for the O K-shell.<sup>5</sup> The theoretical spectrum is reported in Fig. 5, together with the gas phase experimental data. The colored partial contributions of all nonequivalent carbon atoms (Ci) on which the core hole is localized are also highlighted. The calculated site-resolved spectra facilitate the analysis of the transitions and the attribution of the spectral features in terms of the specific carbon atom contributions (see Fig. S9 of the [supplementary material](#)). The detailed analysis of the calculated spectrum can be found in Table S6 of the [supplementary material](#).

Inspection of Fig. 5 reveals the good match between experiment and theory; in particular, the theoretical results correctly describe the energy separation among the peaks as well as the relative intensity distribution. These results therefore prove that the relaxation effects explicitly included in the DFT-TP computational scheme are suitable for correctly describing the C K-edge spectrum of the PBA molecule, also in the presence of important electronic correlation effects related to the aromatic stabilization as well as to charge polarization. As concerns the assignment of the experimental spectral features, the first and most intense peak (peak B) derives from the C1s  $\rightarrow$  LUMO ( $1\pi^*$ ) transitions from all the Ci sites, with the exception of C2, whose  $1\pi^*$  transition gives rise to the lower energy shoulder (peak A at around 284.4 eV), reflecting the increased electrostatic shielding of the C2 atom due to the bond with the boron atom. The  $1\pi^*$  orbital is delocalized on the phenyl ring and to minor extent on the two B-OH moieties, in analogy with the B1s core spectrum attribution. Also, the higher energy shoulder of peak B has a  $\pi^*$  character. Peaks C and D are contributed by several low intensity transitions toward orbitals with significant diffuse C atomic components, while the stronger feature E mainly derives from transitions toward higher

energy  $\pi^*$  orbitals from all the Ci sites except C2 whose corresponding  $\pi^*$  transition falls at lower energy, in the region of the peak D. Transitions to diffuse orbitals also contribute to E and F higher energy features. In summary, the C1s results further support the hypothesis that high-level correlation effects beyond the relaxation contributions specifically influence the B1s core excitations in PBA and PDBA (and likely in other phenyl-boronic acid derivatives) and their inclusion is crucial to obtain a quantitative description of the experimental spectral structures. The present C1s results for PBA represent an important piece of information to approach with more confidence the NEXAFS calculations at the K-edge of the different atoms present in this kind of molecules, also in the perspective of possible applications to more extended systems based on phenyl boronic acids.

## V. CONCLUSIONS

Our study raises a warning for the theoretical interpretation of B1s NEXAFS spectra in phenyl boronic systems of various dimensionality. In this work, we investigated the influence of electron-correlation effects in the theoretical description of B1s core excitations of B(OH)<sub>3</sub>, PBA, and PDBA molecules. High-resolution synchrotron radiation measurements at the B K-edge are also reported and allow a strict comparison between observed and theoretical data. The DFT-TP approach is suitable for correctly describing the B1s NEXAFS spectrum of the simplest B(OH)<sub>3</sub> molecule but fails when applied to the two phenyl boronic molecules, irrespective of the exchange-correlation functional employed. The inadequacy of the one electron scheme shows up in a poor reproduction of the energy separation and the intensity distribution between the first two bands of the experimental spectrum. To obtain a satisfactory agreement with the experiment, correlation effects beyond those described by the coupling of singly excited configurations, as in the TDDFT approach, have to be included in the computational scheme. This was demonstrated by performing B1s core spectra calculations of PBA and PDBA at the MCSCF level of theory which provides a quantitative reproduction of the first two absorption bands. An analysis of the MCSCF wave functions revealed the importance of a doubly excited configuration in the description of the excited state associated with the second peak of the PBA and PDBA B1s spectra, and accounts for the breakdown of both single-particle (DFT) and the 1h-1p configuration interaction (TDDFT) approaches. Concerning the interpretation of the two main peaks of the PBA and PDBA B1s spectrum, it is convenient to rely on the DFT-TP results because the final MOs involved in the excited MCSCF configurations and in the single particle description of the main peaks have the same  $\pi^*$  nature. This is important in the perspective of a B1s NEXAFS spectral interpretation based on the DFT-TP approach of large systems for which the use of explicitly correlated wave function methods has become quickly unfeasible. The theoretical simulation of the C1s NEXAFS spectrum of PBA, for which the high resolution XAS spectrum at the C K-edge was acquired, was performed at the DFT-TP level, and a good match between theory and experiment was obtained, confirming that (i) the relaxation effects explicitly included in the DFT-TP computational scheme are suitable for correctly describing the C K-edge spectrum of this system and (ii) strong electron-correlation effects specifically affect the B1s core excited states of these important



**FIG. 5.** C-K edge NEXAFS spectrum of PBA: experimental spectrum (blue circles) and calculated line-shape (black solid line) with partial Ci contributions (thick colored vertical bars). The chemical structure of PBA is reported in the inset. The calculated spectrum was rigidly shifted on the experimental energy scale by  $-0.45$  eV in order to match the experimental results. Colored vertical dashed lines:  $\Delta$ KS C1s ionization thresholds.

boronic-acid derivatives. We are led to assume that the relaxation of the charge density following the core-hole formation can be less pronounced for the shallower B1s core state, compared to C or N ones, and residual electron correlation effects (quasidegenerate correlation effects) can play a more important role in the description of the core excitation process than in the heavier first row atoms such as C, N, and O. Another aspect, which is however specific to the molecules under study, is connected to their planarity [i.e., the B(OH)<sub>2</sub> moiety is coplanar with the phenyl ring] with consequent additional electron delocalization effects, which are challenging to describe at the TP approximation. In particular, in the case of boron, the effective +0.5 charge used in the standard TP method is unable to correctly describe the extent of electron delocalization upon the formation of the core hole. In this respect, the MCSCF approach is better suited to describe the interaction of the excited B1s electron with the valence charge density (due to a more accurate, correlated description of the core-excited final states beyond 1h-1p).

## SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for further details about calculations, optimized equilibrium geometries, and comparison with the experimental data (Tables S1–S4); full assignment of B1s NEXAFS spectra of B(OH)<sub>3</sub>, PBA, and PDBA (Table S5); DFT and TDDFT results for B1s NEXAFS spectra of B(OH)<sub>3</sub>, PBA and PDBA not included in the manuscript (Figs. S1–S7); and experimental C1s NEXAFS spectrum (Fig. S8) and full assignment of the C1s NEXAFS spectrum of PBA (Fig. S9, Table S6).

## ACKNOWLEDGMENTS

Computational research was supported by Finanziamento per ricerca di ateneo, FRA 2015 and FRA 2016 of the Università degli Studi di Trieste.

The authors acknowledge the Carl Tygger Foundation for making the VG-Scienta SES-200 photoelectron analyzer available at the Gas Phase Photoemission beam line, Elettra, Italy.

We thank Fabio Zuccaro for technical support in all phases of the experiment.

## REFERENCES

- <sup>1</sup>D. G. Hall, *Boronic Acids. Preparation and Applications in Organic Synthesis and Medicine* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG, 2006).
- <sup>2</sup>R. Nishiyabu, Y. Kubo, T. D. James, and J. S. Fossey, *Chem. Commun.* **47**, 1124 (2011); Y. Kubo, R. Nishiyabu, and T. D. James, *ibid.* **51**, 2005 (2014).
- <sup>3</sup>S. Clair, M. Abel, and L. Porte, *Chem. Commun.* **50**, 9627 (2014); N. A. A. Zwanveld, R. Pawlak, M. Abel, D. Catalin, D. Gigmes, D. Bertin, and L. Porte, *J. Am. Chem. Soc.* **130**, 6678 (2008).
- <sup>4</sup>M. Stredansky, A. Sala, T. Fontanot, R. Costantini, C. Africh, G. Comelli, L. Floreano, A. Morgante, and A. Cossaro, *Chem. Commun.* **54**, 3971 (2018).
- <sup>5</sup>D. Toffoli, M. Stredansky, Z. Feng, G. Balducci, S. Furlan, M. Stener, H. Ustunel, D. Cvetko, G. Kladnik, A. Morgante, A. Verdini, C. Dri, G. Comelli, G. Fronzoni, and A. Cossaro, *Chem. Sci.* **8**, 3789 (2017).
- <sup>6</sup>J. C. Slater, *Adv. Quantum Chem.* **6**, 1 (1972); B. O. Roos, *Adv. Chem. Phys.* **69**, 399 (1987).
- <sup>7</sup>L. Triguero, L. G. M. Pettersson, and H. Ågren, *Phys. Rev. B* **58**, 8097 (1998).
- <sup>8</sup>H. Lischka, D. Nachtigallova, A. J. A. Aquino, P. G. Szalay, F. Plasser, F. B. C. Machado, and M. Barbatti, *Chem. Rev.* **118**, 7293 (2018).
- <sup>9</sup>K. C. Prince, R. R. Blyth, R. Delaunay, M. Zitnik, J. Krempasky, J. Slezak, R. Camilloni, L. Avaldi, M. Coreno, G. Stefani, C. Furlani, M. de Simone, and S. Stranges, *J. Synchrotron Radiat.* **5**, 565 (1998).
- <sup>10</sup>R. N. Sodhi and C. E. Brion, *J. Electron Spectrosc. Relat. Phenom.* **34**, 363 (1984).
- <sup>11</sup>N. Martensson, P. Baltzer, P. A. Brühwiler, J.-O. Forsell, A. Nilsson, A. Stenborg, and B. Wannberg, *J. Electron Spectrosc. Relat. Phenom.* **70**, 117 (1994).
- <sup>12</sup>L. Pettersson, J. Nordgren, L. Selander, K. Siegbahn, and H. Ågren, *J. Electron Spectrosc. Relat. Phenom.* **27**, 29 (1982).
- <sup>13</sup>V. Myrseth, J. D. Bozek, E. Kukk, L. J. Sæthre, and T. D. Thomas, *J. Electron Spectrosc. Relat. Phenom.* **122**, 57 (2002).
- <sup>14</sup>S. H. Vosko, L. Wilk, and M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).
- <sup>15</sup>E. J. Baerends, D. E. Ellis, and P. Roos, *Chem. Phys.* **2**, 41 (1973); C. Fonseca Guerra, J. G. Snijders, G. te Velde, and E. J. Baerends, *Theor. Chem. Acc.* **99**, 391 (1998).
- <sup>16</sup>N. Z. Rao, J. D. Larkin, and C. W. Bock, *Struct. Chem.* **28**, 945 (2017).
- <sup>17</sup>R. Pawlak, L. Nony, F. Bocquet, V. Oison, M. Sassi, J.-M. Debierre, C. Loppacher, and L. Porte, *J. Phys. Chem. C* **114**, 9290 (2010).
- <sup>18</sup>J. P. Perdew, *Phys. Rev. B* **33**, 8822 (1986); A. Bonasera, *Phys. Rev. C* **34**, 740 (1986).
- <sup>19</sup>P. Norman and A. Dreuw, *Chem. Rev.* **118**, 7208 (2018).
- <sup>20</sup>A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993); C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988); P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).
- <sup>21</sup>Y. Zhao and D. G. Truhlar, *J. Chem. Phys.* **125**, 194101 (2006); *Theor. Chem. Acc.* **120**, 215 (2008).
- <sup>22</sup>E. K. U. Gross and W. Kohn, *Adv. Quantum Chem.* **21**, 255 (1990).
- <sup>23</sup>M. Stener, G. Fronzoni, and M. de Simone, *Chem. Phys. Lett.* **373**, 115 (2003).
- <sup>24</sup>R. van Leeuwen and E. J. Baerends, *Phys. Rev. A* **49**, 2421 (1994).
- <sup>25</sup>M. Ernzerhof and G. Scuseria, *J. Chem. Phys.* **110**, 5029 (1999); C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).
- <sup>26</sup>B. O. Roos, *Adv. Chem. Phys.* **69**, 399 (1987).
- <sup>27</sup>H.-J. Werner and P. J. Knowles, *J. Chem. Phys.* **82**, 5053 (1985); P. J. Knowles and H.-J. Werner, *Chem. Phys. Lett.* **115**, 259 (1985).
- <sup>28</sup>H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang, and A. Wolf, *MOLPRO*, version 2010.1, a package of *ab initio* programs, 2010, see <http://www.molpro.net>.
- <sup>29</sup>M. Alagia, E. Bodo, P. Declava, S. Falcinelli, A. Ponzi, R. Richterf, and S. Stranges, *Phys. Chem. Chem. Phys.* **15**, 1310 (2013).
- <sup>30</sup>P. Elliott, S. Goldson, C. Canahui, and N. T. Maitra, *Chem. Phys.* **391**, 110 (2011).
- <sup>31</sup>N. T. Maitra, *J. Chem. Phys.* **144**, 220901 (2016).
- <sup>32</sup>M. Huix-Rotlant, A. Ipatov, A. Rubio, and M. E. Casida, *Chem. Phys.* **391**, 120 (2011).

# Correlation effects in B1s core-excited states of boronic-acid derivatives: an experimental and computational study

Daniele Toffoli<sup>a</sup>, Aurora Ponzi<sup>b</sup>, Elisa Bernes<sup>a</sup>, Monica de Simone<sup>c</sup>, Cesare Grazioli<sup>c</sup>, Marcello Coreno<sup>d</sup>, Matus Stredansky<sup>c,e</sup>, Albano Cossaro<sup>c</sup>, Giovanna Fronzoni<sup>a,\*</sup>

<sup>a</sup> Department of Chemical and Pharmaceutical Sciences, University of Trieste, via L. Giorgieri 1, 34127, Trieste (Italy). E-mail: [fronzoni@units.it](mailto:fronzoni@units.it)

<sup>b</sup> Ruđer Bošković Institute, Bijenička cesta 54, 10000, Zagreb (Croatia)

<sup>c</sup> CNR-IOM Laboratorio TASC, Basovizza SS-14, km 163.5, 34149 Trieste (Italy)

<sup>d</sup> CNR-ISM (LD2 Unit) Basovizza SS-14, Km 163,5 34149 Trieste (Italy)

<sup>e</sup> Department of Physics, University of Trieste, via A. Valerio 2, 34127 Trieste (Italy)

## Supplementary Material

### Details on the Theoretical Methods

In the DFT-TP approach the ionization energy threshold is obtained from the KS eigenvalue:

$$E_{ion} = -\varepsilon_i^{TP} \quad (1)$$

The TP approach leads to a less attractive potential and the absolute transition energies are generally too large. In order to correct the NEXAFS energies,  $\Delta$ KS ( $\Delta$ SCF Kohn-Sham) calculations of the ionization energies have been performed, allowing a full relaxation of the ionized core hole. The energy of the  $1s^{-1}$  ionic state is obtained from an unrestricted KS calculation. The DFT-TP excitation energies are then shifted with respect to the  $\Delta$ IP value corresponding to  $[\varepsilon_{1s}^{TP} - \Delta KS]$ .

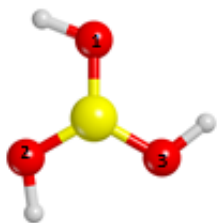
## Results

**Table S1.** Calculated equilibrium geometry of  $B(OH)_3$ , PBA and PDBA

ATOM	X(Å)	Y(Å)	Z(Å)
<b>B(OH)<sub>3</sub></b>			
B	0,000054	0,000013	0,00000
O	0,000249	1,363462	0,00000
O	-1,181008	-0,681344	0,00000
O	1,180702	-0,682003	0,00000
H	-0,904416	1,726263	0,00000
H	-1,042994	-1,646220	0,00000
H	1,947623	-0,080410	0,00000
<b>PBA</b>			
B	0,0000000	0,0096633	-1,70810337
C	0,0000000	0,0223251	-0,14043806
C	0,0000000	1,2117228	0,60154157
C	0,0000000	1,1983228	1,99392115
C	0,0000000	-0,0175929	2,67435209
C	0,0000000	-1,2128468	1,95691803
C	0,0000000	-1,1900545	0,56527269
H	0,0000000	-2,1177759	0,00890669
H	0,0000000	-2,1580718	2,48272471
H	0,0000000	-0,0329436	3,75588480
H	0,0000000	2,1285110	2,54569437
H	0,0000000	2,1735712	0,09997797
O	0,0000000	1,1417296	-2,48725168
H	0,0000000	1,9485114	-1,96808006
O	0,0000000	-1,2020278	-2,34386601
H	0,0000000	-1,0876129	-3,29963972
<b>PDBA</b>			
C	0,660622	1,211186	0,000000
C	1,407880	0,034090	0,000000
B	2,957905	0,041889	0,000000
C	0,721823	-1,180660	0,000000
H	1,164825	2,187710	0,000000
C	-0,720926	1,183225	0,000000
C	-0,659753	-1,208611	0,000000
C	-1,407005	-0,031532	0,000000
B	-2,957088	-0,039384	0,000000
H	1,301911	-2,108650	0,000000
O	3,702710	1,184288	0,000000
O	3,614613	-1,148585	0,000000
O	-3,613831	1,151048	0,000000
O	-3,701716	-1,181897	0,000000
H	4,582176	-1,020271	0,000000

H	-4,581426	1,022855	0,000000
H	-1,163805	-2,185223	0,000000
H	-1,301037	2,111207	0,000000
H	3,147750	1,985732	0,000000
H	-3,146500	-1,983197	0,000000

**TABLE S2** - Comparison between theoretical and experimental geometrical parameters for  $B(OH)_3$ . Bond lengths are expressed in Å, angles in degrees.

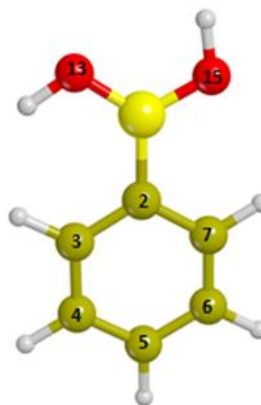


<b>Bond lengths</b>	Calculated	Experimental <sup>a</sup>
B-O1	1.3634	1.3610
B-O2	1.3635	1.3640
B-O3	1.3635	1.3630
<b>Angles</b>	Calculated	Experimental <sup>a</sup>
O1-B-O2	119.989	119.900
O1-B-O3	120.005	120.400
O2-B-O3	120.006	119.600

<sup>a</sup>R.R.Shuvalov, P.C.Burns, *Acta Crystallographica, Section C: Crystal Structure Communications*, 2003, 59, i47-i49

**TABLE S3** - Comparison between theoretical and experimental geometrical parameters for PBA. Bond lengths are expressed in Å, angles in degrees.

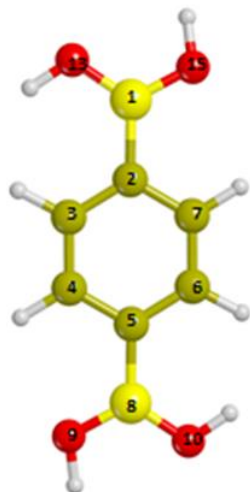
Bond lengths	Calculated	Experimental <sup>a</sup>
C2-B	1.5677	1.5660
B-O13	1.3743	1.3750
B-O15	1.3684	1.3668
C2=C3	1.4019	1.4025
C2-C7	1.4028	1.4070
C3-C4	1.3924	1.3920
C4=C5	1.3934	1.3860
C5-C6	1.3940	1.3760
C6=C7	1.3918	1.3915
Angles	Calculated	Experimental <sup>b</sup>
C-B-O13	124.075	120.150
C-B-O15	118.148	119.200
O13-B-O15	117.777	115.950
B-C2=C3	122.420	121.950
B-C2-C7	119.740	120.600
C2=C3-C4	121.406	121.150
C2-C7=C6	121.141	121.150
C3=C2-C7	117.840	117.450
C3-C4=C5	119.783	120.000
C4=C5-C6	119.795	121.500



<sup>a</sup>S.J.Rettig, J.Trotter, *Can.J.Chem.*, 1977, 55, 3071-3075



**TABLE S4** - Comparison between theoretical and experimental geometrical parameters for PDBA. Bond lengths are expressed in Å, angles in degrees.



Bond lengths	Calculated	Experimental <sup>a</sup>
C2-B1	1.5501	1.5640
C5-B8	1.5500	
B1-O13	1.3637	1.3600
B1-O15	1.3596	1.3490
B8-O10	1.3637	1.3600
B8-O9	1.3596	1.3490
C2=C3	1.3942	1.3840
C2-C7	1.3951	1.3900
C3-C4	1.3816	1.3810
C4=C5	1.3951	1.3900
C5-C6	1.3943	1.3840
C6=C7	1.3818	1.3810
Angles	Calculated	Experimental <sup>c</sup>
C2-B1-O13	123.384	122.620
C2-B1-O15	118.595	119.670
C5-B8-O10	123.391	122.620
C5-B8-O9	118.594	119.670
O13-B1-O15	118.021	117.700
O10-B8-O9	118.014	
B1-C2=C3	122.119	121.920
B1-C2-C7	119.747	120.740
C2=C3-C4	121.250	121.440
C2-C7=C6	120.617	121.230
C3=C2-C7	118.134	117.330
C3-C4=C5	120.616	121.230
C4=C5-C6	118.135	117.330
C5-C6=C7	121.249	121.440

<sup>a</sup>P. Rodríguez-Cuamatzi, G. Vargas-Díaz, T. Maris, J.D. Wuest, H. Höpfl, *Acta Crystallographica, Section E: Structure Reports Online*, 2004, 60 (8), o1316- o1318

The following Table S5 reports the calculated B1s excitation energies E(eV) and oscillator strengths (f) useful to assign the experimental features on the basis of the DFT-TP results for the B(OH)<sub>3</sub>, PBA and PDBA obtained with the PW86 xc potential. Figures are reported in the main text (upper panel of : Figure 1 for the B(OH)<sub>3</sub>, Figure 2 for PBA and Figure 4 for PDBA). Due to the low cross section of diffuse Rydberg MOs and the low density of the samples in gas phase, these high energy states results hidden in the experimental noise.

**Table S5.** B1s excitation energies E(eV) and oscillator strengths (f) for the main transitions below edge of B(OH)<sub>3</sub>, PBA and PDBA calculated with DFT-TP (PW86 xc potential) approach. Experimental energies are reported in the last column

E(eV) <sup>a</sup>	E (eV) shifted <sup>b</sup>	f <sup>c</sup>	Assignment	Experiment (eV) <sup>d</sup>
<b>B(OH)<sub>3</sub></b>				
194.74	194.27	0.088	1π* [main contributions: 2p <sub>z</sub> B ,O ]	194.27(2)
				193.85(6) 194.85(2) vibrational features
196.310	195.840	0.0021	diffuse Rydberg MOs	195.74(4)
196.313	195.843	0.0021		
197.22	196.75	0.0017	diffuse Rydberg MOs	197.02(9)
197.60	197.13	0.0020		
197.61	197.14	0.0024		
<b>PBA</b>				
191.99	192.03	0.0408	1π* [main contributions: 2p <sub>z</sub> C, B ; minor 2p <sub>z</sub> O ]	192.03(1)
194.57	194.61	0.0369	2π* [main contributions: 2p <sub>z</sub> C, B ; minor 2p <sub>z</sub> O ]	194.20(1)
194.77	194.81	0.0020	diffuse Rydberg MOs	
195.65	195.69	0.0022		
195.78	195.82	0.0016		
195.94(1)				
<b>PDBA</b>				
191.81	192.00	0.0691	1π* [main contributions: 2p <sub>z</sub> C, B ; minor 2p <sub>z</sub> O ]	192.00(1)
194.07	194.26	0.0762	π* [main contributions: 2p <sub>z</sub> C, B ; minor 2p <sub>z</sub> O ]	193.97(1)

194.78	194.97	0.0036	diffuse Rydberg MOs	195.66(1)
196.17	196.36	0.0050		
196.29	196.48	0.0021		
196.42	196.61	0.0069	mixed Rydberg/valence $\pi^*$ MOs	
196.60	196.79	0.0116		

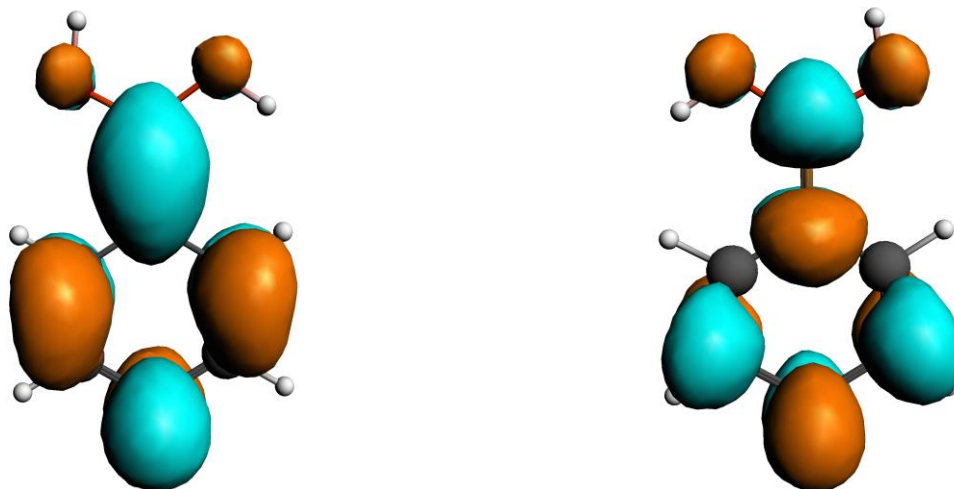
<sup>a</sup>Calculated excitation energies. Calculated IP- $\Delta$ SCF: 198.58 eV for B(OH)<sub>3</sub>; 196.93 eV for PBA; 196.94 eV for PDBA.

<sup>b</sup>Calculated excitation energies shifted by -0.47 eV (B(OH)<sub>3</sub>), 0.04 eV (PBA), 0.19 eV (PDBA) to match the first experimental peak.

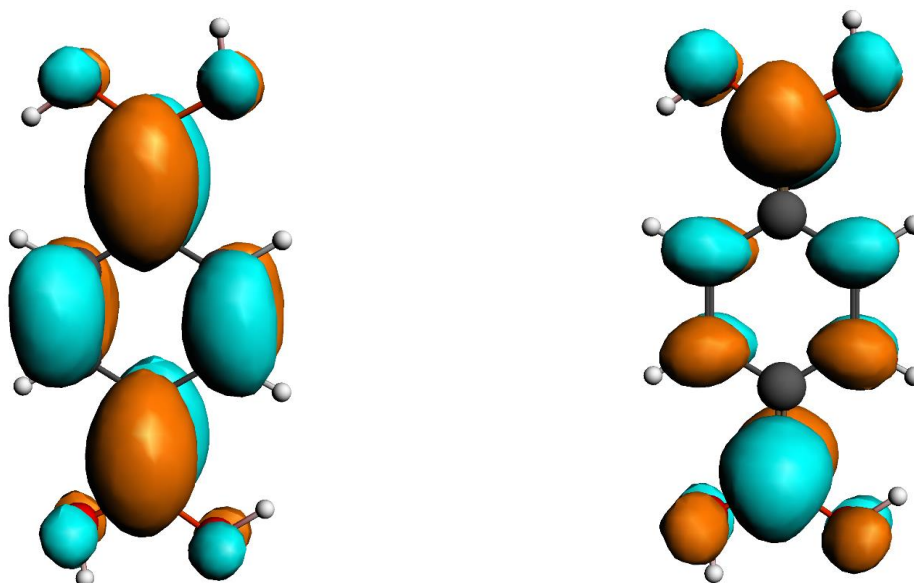
<sup>c</sup>Only transitions with oscillator strength  $f \geq 0.0015$  are reported.

<sup>d</sup>The experimental energies have been extracted from measured data, fitting the B K-edge spectrum with a Gaussian curve for each main visible structure in the spectrum.

a)

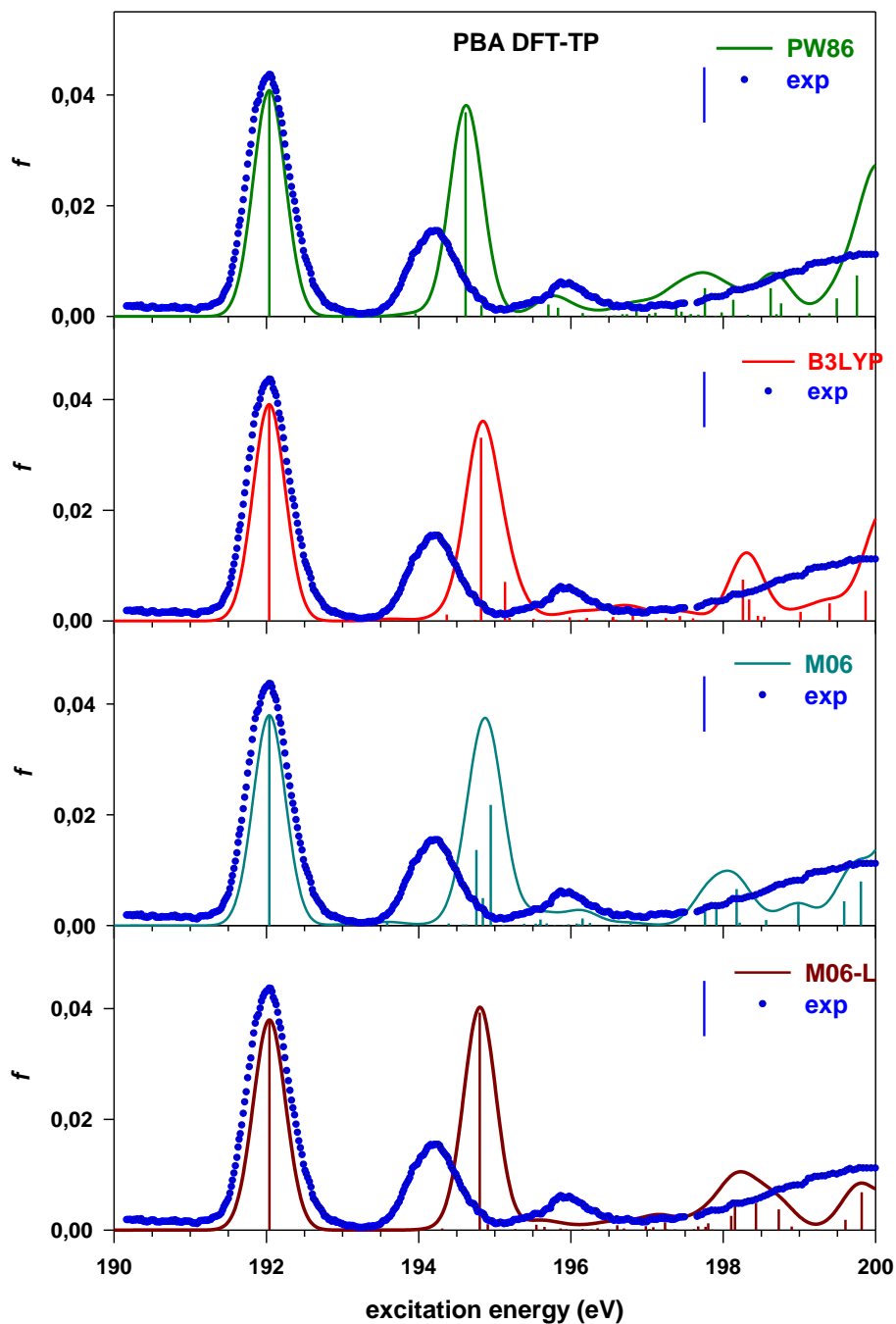


b)

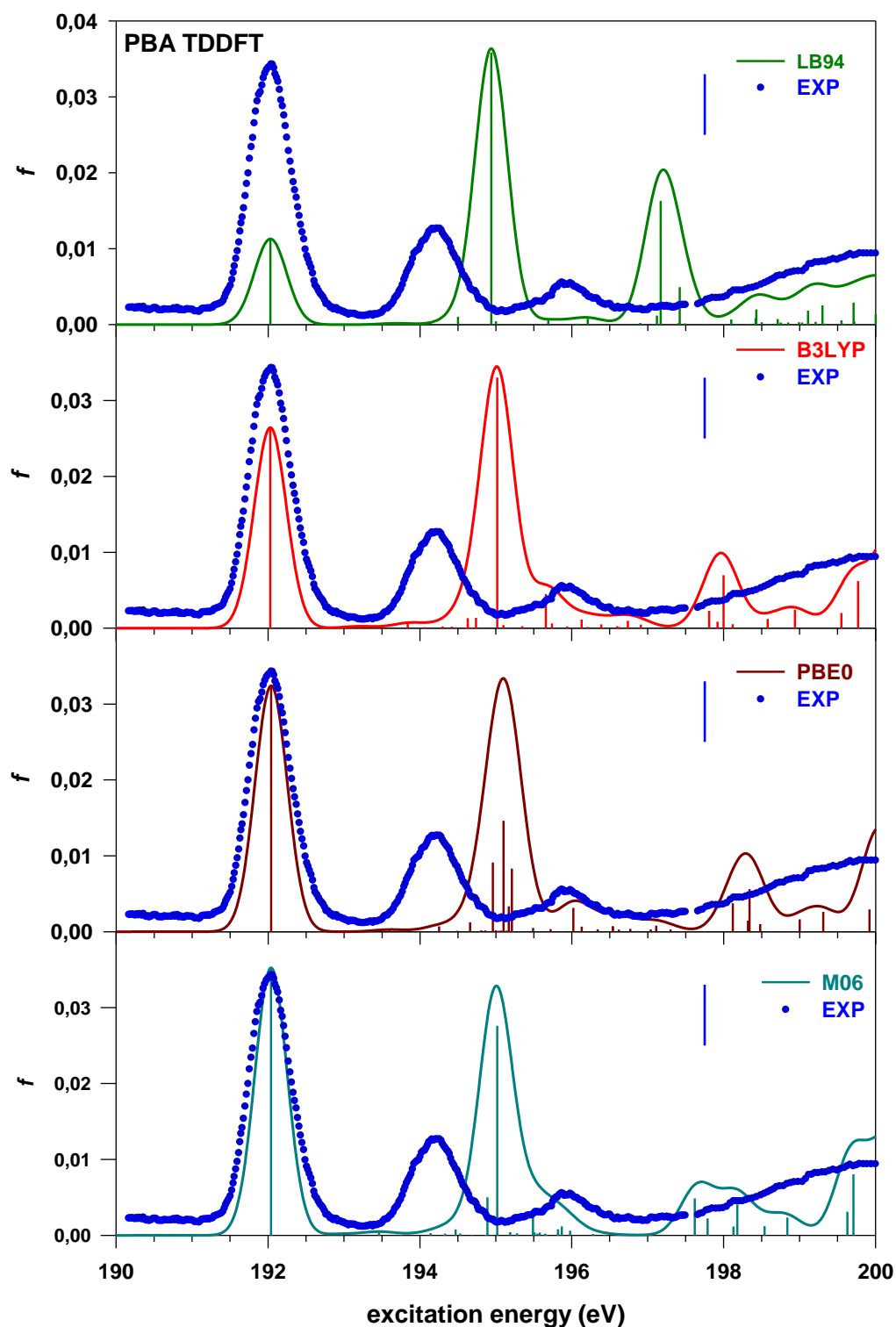


**Fig. S1.** Plot of the calculated DFT-TP (PW86 xc potential) unoccupied MOs involved in the two main  $B1s$  transitions of PBA and PDBA spectra. Panel a: LUMO (left) and LUMO+3 (right) MOs of PBA; panel b: LUMO (left) and LUMO+3 (right) MOs of PDBA. Displayed isosurface corresponds to  $\pm 0.03 e^{1/2} a_0^{-3/2}$  value.

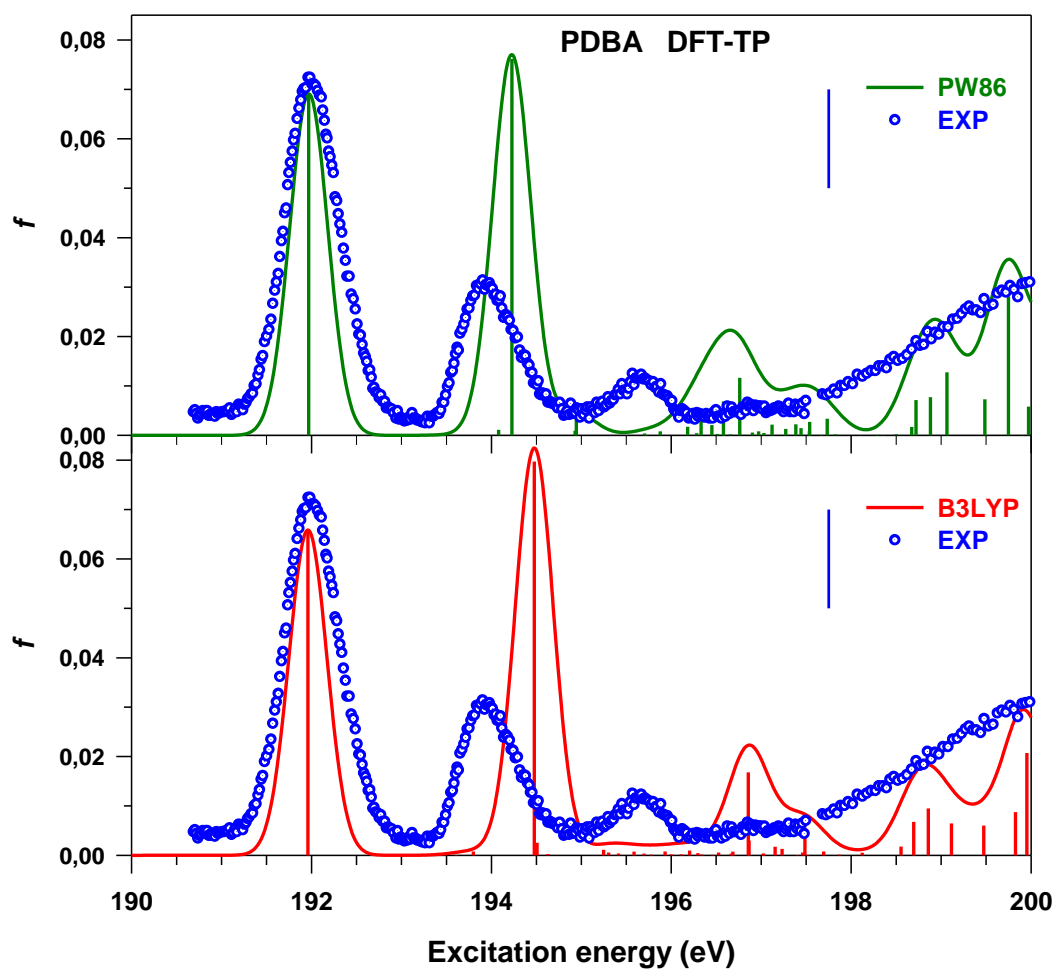
## B1s NEXAFS spectra



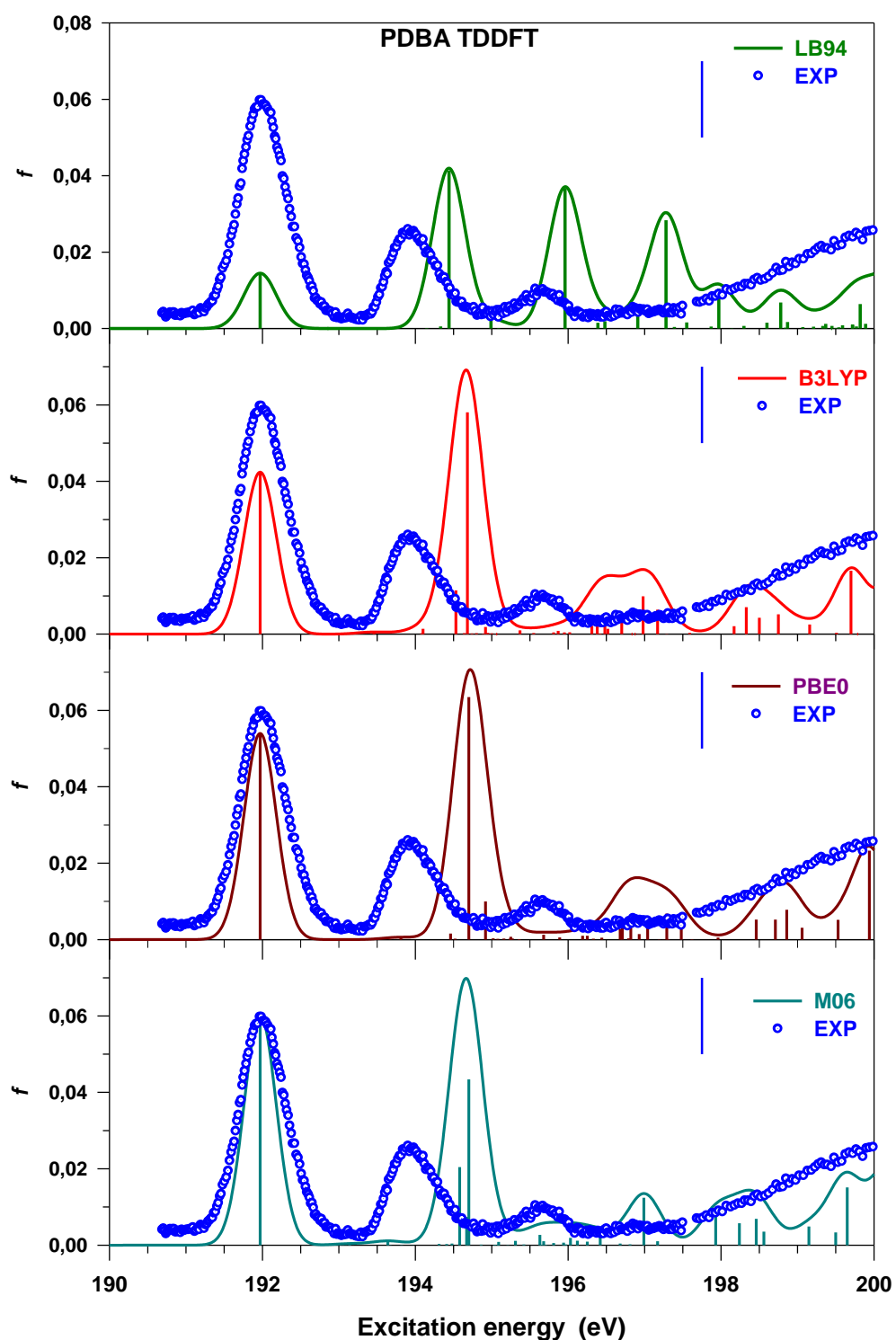
**Fig. S2.** Calculated B1s NEXAFS spectrum of PBA at DFT-TP level employing the exchange correlation potential indicated by the legend. The experimental spectrum is reported for comparison. A rigid shift has been applied to the calculated B1s profiles to align the energy position of the first absorption peak to the experimental one (+0.04 eV for PW86, -0.84 eV for B3LYP, -0.48 eV for M06 and +0.33 eV for M06-L). Vertical blue lines: IP (B1s) experimental value.



**Fig. S3.** Calculated B1s NEXAFS spectrum of PBA at TDDFT level employing exchange correlation potentials with increasing fractions of Hartree–Fock (HF) exchange. The experimental spectrum is reported for comparison. A rigid shift has been applied to the calculated B1s profiles to align the energy position of the first absorption peak to the experimental one ( 3.29 eV for LB94; 10.16 eV for B3LYP; 9.60 eV for PBE0; 9.23 eV for M06). Vertical blue lines: IP (B1s) experimental value.

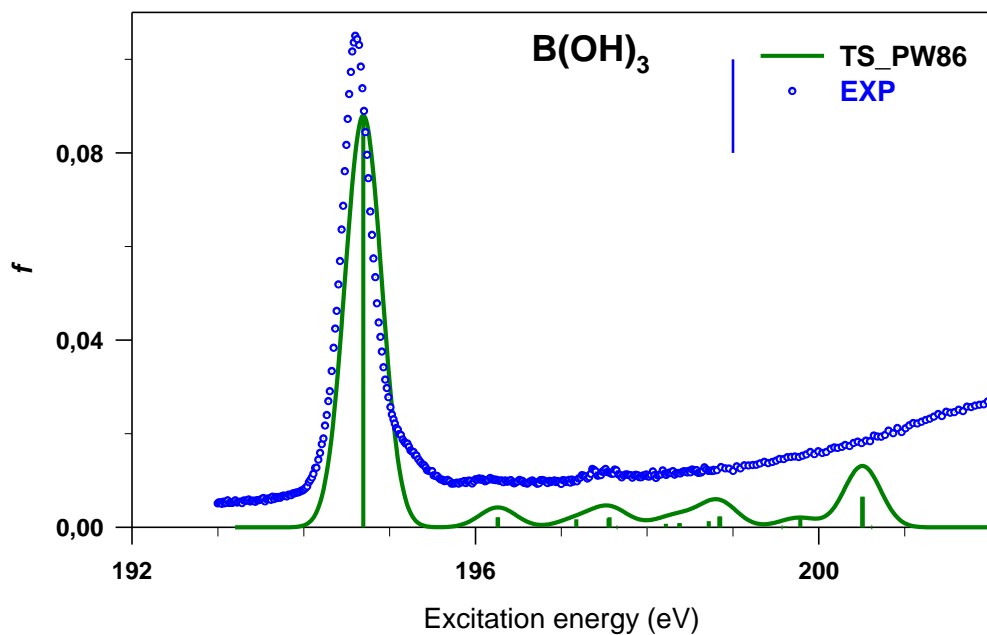


**Fig. S4.** Calculated NEXAFS  $B(1s)$  of PDBA at DFT-TP level employing the exchange correlation potential indicated by the legend. The experimental spectrum is reported for comparison. A rigid shift has been applied to the calculated  $B1s$  profiles to align the energy position of the first absorption peak to the experimental one (0.19 eV for PW86, -0.64 eV for B3LYP). Vertical blue lines: IP ( $B1s$ ) experimental value.

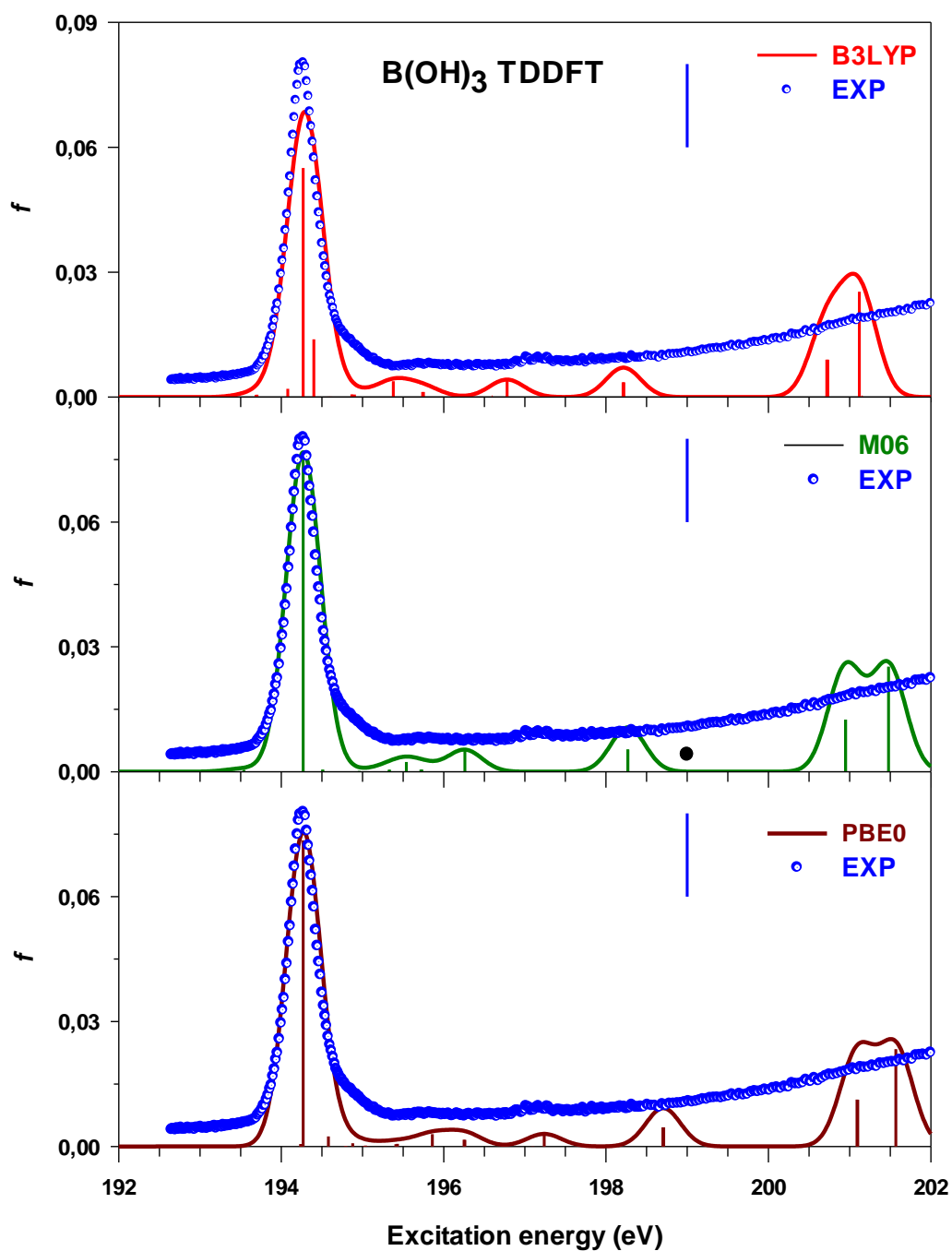


**Fig. S5.** Calculated NEXAFS  $B(1s)$  of PDBA at TDDFT level employing exchange correlation potentials with increasing fractions of Hartree–Fock (HF) exchange. The experimental spectrum is reported for comparison. A rigid shift has been applied to the calculated  $B1s$  profiles to align the energy position of the first absorption peak to the experimental one (3.74 eV for LB94; 10.41 eV for B3LYP; 9.82 eV for PBE0; 9.42 eV for M06). Vertical blue lines: IP ( $B1s$ ) experimental value.



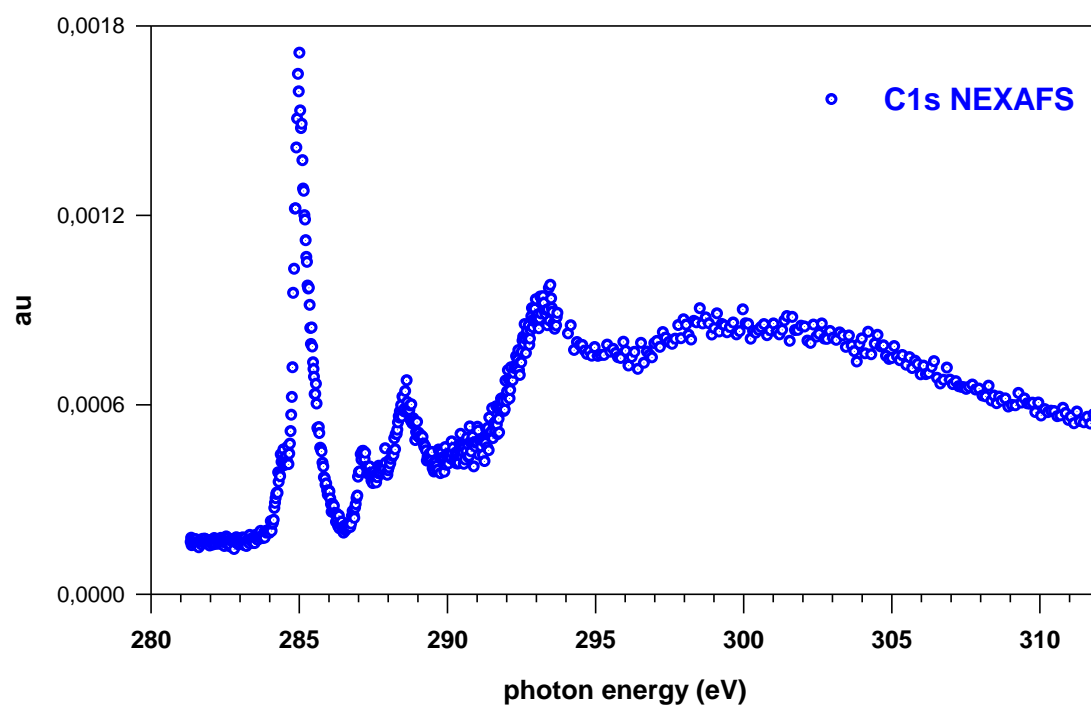


**Fig. S6.** Calculated NEXAFS B(1s) of B(OH)<sub>3</sub> at DFT-TP level with PW86 exchange correlation potential. The experimental spectrum is reported for comparison. A rigid shift of 0.47 eV has been applied to the calculated B1s profiles to align the energy position of the first absorption peak to the experimental one. Vertical blue lines: IP (B1s) experimental value.

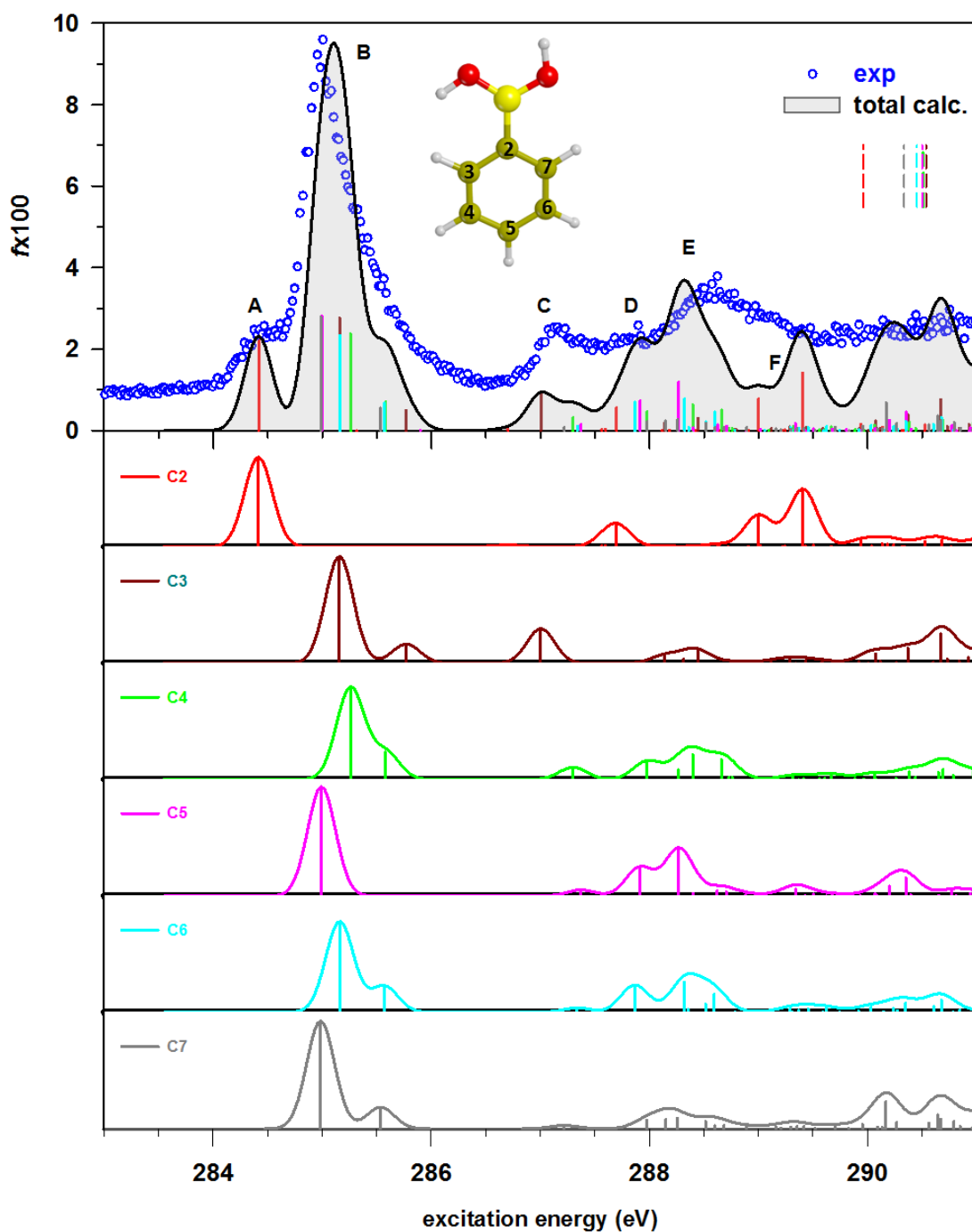


**Fig. S7.** Calculated NEXAFS B(1s) of B(OH)<sub>3</sub> at TDDFT level employing a selection of exchange correlation potentials. The experimental spectrum is reported for comparison. A rigid shift has been applied to the calculated B1s profiles to align the energy position of the first absorption peak to the experimental one (9.16 eV for B3LYP; 8.62 eV for PBE0; 8.81 eV for M06). Vertical blue lines: IP (B1s) experimental value.

### C1s NEXAFS spectra



**Fig. S8.** Experimental C K-edge NEXAFS spectrum of PBA



**Fig. S9.** C K-edge NEXAFS spectrum of PBA. Upper panel: experimental spectrum (blue circles) and calculated lineshape (black solid line) with partial  $C_i$  contributions (thick colored vertical bars) calculated at DFT-TP level employing the PW86 exchange correlation potential. The chemical structure of PBA is reported in the inset. Lower panels: Partial  $C_{1s}$  DFT-TP calculated spectra. All the calculated spectra have been rigidly shifted on the experimental energy scale by  $-0.45$  eV in order to match the experimental results. Colored vertical dashed lines:  $\Delta K_S$   $C_{1s}$  ionization thresholds (C2: 289.96 eV; C3: 290.54 eV; C4: 290.51 eV; C5: 290.50 eV; C6: 290.45 eV; C7: 290.34 eV).

**Table S6.** Peak assignments for the C K-edge NEXAFS spectrum of PBA calculated at DFT-TP level employing the PW86 exchange correlation potential. Experimental energies are reported in the last column.

Peak	Calculated			Assignment	Experiment
	Site	E(eV) shifted <sup>a</sup>	f x10 <sup>2</sup> <sup>b</sup>		
<b>A</b>	C <sub>2</sub>	284.42	2.32	1 $\pi^*$ (LUMO)	284.38+/-0.02
<b>B</b>	C <sub>7</sub>	284.98	2.78	1 $\pi^*$ (LUMO)	284.98+/-0.01
	C <sub>5</sub>	284.99	2.82		
	C <sub>3</sub>	285.15	2.76		
	C <sub>6</sub>	285.16	2.34		
	C <sub>4</sub>	285.26	2.37	2 $\pi^*$ (LUMO+1)	285.25+/-0.04
	C <sub>7</sub>	285.53	0.55		
	C <sub>6</sub>	285.57	0.68		
	C <sub>4</sub>	285.58	0.71		
<b>C</b>	C <sub>3</sub>	287.00	0.89	mixed Rydberg/ valence $\sigma^*$ (O-H)	287.17+/-0.04
	C <sub>4</sub>	287.30	0.32		
	C <sub>5</sub>	287.37	0.16		
<b>D</b>	C <sub>2</sub>	287.70	0.56	$\pi^*$	287.71+/-0.05
	C <sub>6</sub>	287.87	0.70	mixed Rydberg/ valence $\sigma^*$ (O-H)	
	C <sub>5</sub>	287.91	0.74		
	C <sub>7</sub>	287.97	0.22		
	C <sub>4</sub>	287.98	0.47		
<b>E</b>	C <sub>3</sub>	288.13	0.19	Rydberg	288.66+/-0.07
	C <sub>7</sub>	288.15	0.24		
	C <sub>4</sub>	288.27	0.25		
	C <sub>7</sub>	288.26	0.28		
	C <sub>5</sub>	288.27	1.19		

	C <sub>6</sub>	288.32	0.78	$\pi^*$	
	C <sub>4</sub>	288.40	0.64		
	C <sub>3</sub>	288.44	0.31		
	C <sub>6</sub>	288.51	0.21	Rydberg	
	C <sub>7</sub>	288.52	0.19		
	C <sub>6</sub>	288.60	0.45		
	C <sub>4</sub>	288.66	0.51		
	C <sub>5</sub>	289.34	0.18		
<b>F</b>	C <sub>2</sub>	289.00	0.78	mixed Rydberg/ valence $\pi^*$	
	C <sub>2</sub>	289.41	1.42		

<sup>a</sup>Calculated excitation energies shifted by - 0.45 eV to match the first experimental peak

<sup>b</sup>Only transitions with oscillator strength  $f \times 10^2 \geq 0.15$  are reported.

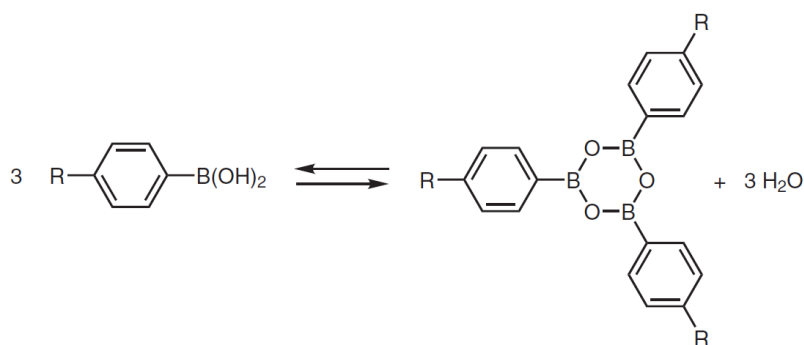
<sup>c</sup>Experimental values have been extracted from measured data, fitting the C K-edge spectrum with a Gaussian curve for each main visible structure (peaks A-E) in the spectrum.

# A NEXAFS Experimental and Theoretical Study on the Nature of the Au-B Interaction in Boroxine-containing Frameworks Deposited on the Au(111) Surface

## Introduction

In the last decades, great advances in the field of organic electronics have occurred, especially due to the related low environmental impact and costs. Organic compounds are considered as the building blocks of living systems as well as promising candidates for electronic and optoelectronic devices thanks to their interesting physical and chemical properties. Furthermore, by using organic films in electronics, it is possible to obtain devices with peculiar characteristics, such as the mechanical flexibility. However, the efficiency of the organic-based devices is lower than that achievable by traditional silicon-based technology. A further limitation in the application of the organic technology is represented by the electronic transport properties at the interface between the organic film and the electrodes of the device<sup>1,2</sup>. The control on the morphology of the interface between the organic thin film and the metallic electrode during the growth of the film is fundamental to obtain a system with good electronic transport properties<sup>3,4</sup>.

The structural and chemical properties of electrodes can be tailored by templating the metallic substrate through the interposition of an interface film. One successful templating strategy is represented by the synthesis of 2D Covalent Organic Frameworks (COFs) via a *guest-host* approach, in which an organic framework templates the surface and hosts the organic guest molecules; the guest-host recognition occurs when the shape of the nanometric cavities of the template exactly matches that of the guest molecules. 2D COFs are an emerging class of porous crystalline materials composed of covalently bonded organic molecules on a substrate. Their formation can occur either in solution or under vacuum, through a wide variety of chemical reactions (*e.g.*, formation of imines, Ullman reaction, polyester condensation, etc.)<sup>5-8</sup>, leading to a great variability of chemical properties. In particular, the boronate esterification is well exploited to build molecular nanoarchitectures, thanks to the fast and stable formation of the B–O covalent bond.<sup>9</sup> Another powerful synthetic strategy for 2D COFs is represented by the self-condensation of boronic acids, which leads to the formation of boroxine rings (B<sub>3</sub>O<sub>3</sub>): as shown in Figure 1, a boroxine ring is formed by the condensation of three boronic acid groups, accompanied by the release of three water molecules.



**Figure 1** – Formation of the boroxine through the condensation of three boronic acid groups.

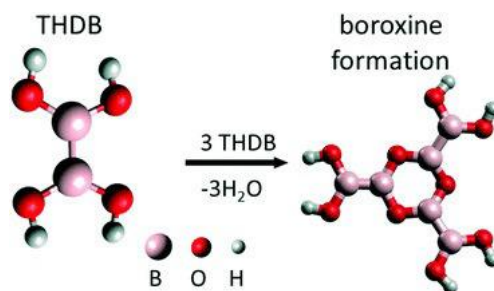
Boroxine-based COFs display interesting properties in terms of crystallinity and thermal stability, and show a better response to the mechanical strain than other 2D materials, such as graphene or boron nitride<sup>10</sup>; this makes them particularly suitable for flexible electronics, and good candidates as templates for the growth of complex organic architectures, although they cannot be directly used as electrodes. It appears therefore important to evaluate if they can also tailor the electronic properties of the electrodes, and, in this respect, a detailed knowledge on the boroxine–metal interface becomes crucial.

Recently, a study on the assembly of phenylboronic acid on the Au(111) surface has shown the formation of triphenylboroxines through spontaneous condensation of trimers of molecules, and revealed that the boroxine group, even on weakly reactive surfaces, represents a promising building block for the functionalization of electrodes in organic electronics. Furthermore, 2D boroxine COFs can be employed to tailor the electronic transport between the guest molecules and the substrate.<sup>11</sup> Another important result of this study concerns the existence of an ultra-fast charge delocalization channel at the boroxine-gold interface<sup>12</sup> due to the presence of an unoccupied electronic state, localized on both the O atoms of the adsorbed triphenylboroxine and the Au atoms of the topmost layer. The channel can be therefore exploited to boost *out-of-plane* charge transport from the boroxine to the metal, which is particularly important to achieve an effective implementation of these systems in organic electronics. On the other hand, boroxine-based COFs exhibit only limited *in-plane* charge carrier mobility, due to the low electronic coupling between the boroxine and phenyl rings usually present in the COF precursors<sup>13</sup>, which hampers the formation of dispersive electronic band structures close to the Fermi level.

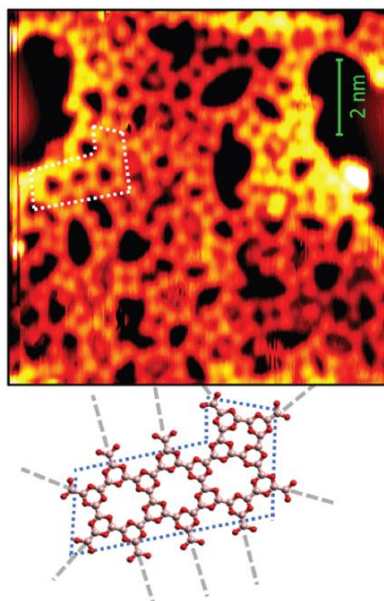
To evaluate whether this limitation in the intrinsic charge transport can be overcome or not, a group of experimental researchers, afferent to the ALOISA beamline of the Elettra Synchrotron in Trieste, has recently synthesized, through *physical vapor deposition*, a phenyl-free 2D framework



based on the boronic acid condensation by employing tetrahydroxydiboron (THDB,  $\text{H}_4\text{B}_2\text{O}_4$ , left side of Figure 2) as precursor<sup>14</sup>. As shown in Figure 2, the condensation of three THDB molecules leads to the formation of a boroxine ring-containing molecule.

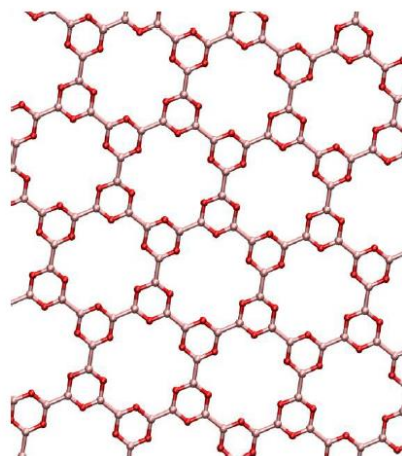


**Figure 2** – Formation of a boroxine through the condensation of three THDB molecules (adapted from [14]). In this work, the formation of the 2D COF has been followed by *X-Ray Photoemission Spectroscopy* (XPS) and *Near Edge X-Ray Absorption Fine Structure Spectroscopy* (NEXAFS), whereas its morphology has been investigated by means of *Scanning Tunnelling Spectroscopy* (STM). In particular, the acquired STM images have revealed a striped phase at low temperatures (180 K), suggesting a head-to-tail hydrogen bonding determined by the boronic–boronic affinity between adjacent intact THDB molecules, while, in between 300 and 450 K, a vitreous 2D network with small bright spots, attributable to the boroxine rings, has been observed. Moreover, a morphological disorder has been detected, with a proliferation of defects in different geometries of connectivity between the boroxine centers, such as 4-fold, 5-fold, and 7-fold motifs; these structural defects induce a strain, limiting the charge delocalization in the system, and preventing the formation of an ideal crystalline structure. A possible structural model of the condensed framework could be formed by three 6-fold and one 4-fold symmetry boroxine motifs, as indicated in the lower panel of Figure 3.



**Figure 3** – Upper panel: STM image of the condensed monolayer ( $T_{\text{sample}} = 77 \text{ K}$ ,  $V = 1 \text{ V}$ ,  $I = 1.5 \text{ nA}$ ). Lower panel: Possible structural model of the motif that could be present in the 2D COF, formed by three 6-fold and one 4-fold symmetry boroxine condensed assemblies molecules (adapted from [14]).

We point out that the 6-fold motif can be considered as the ideal building block for the formation of a periodic 2D boronic network, as reported in Figure 4.

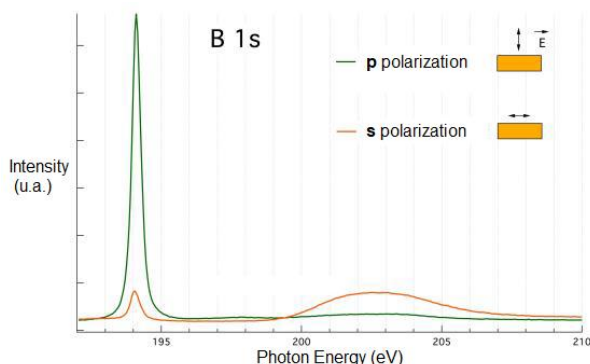


**Figure 4** – Representation of a periodic 2D boronic network.

Despite the morphological disorder described above, *Angle Resolved Ultraviolet Photoemission Spectroscopy* (ARUPS) measurements have revealed the presence of delocalized electronic states, which may give good internal charge transport properties to the system, making it a suitable hybrid interface for the design of electrodes.

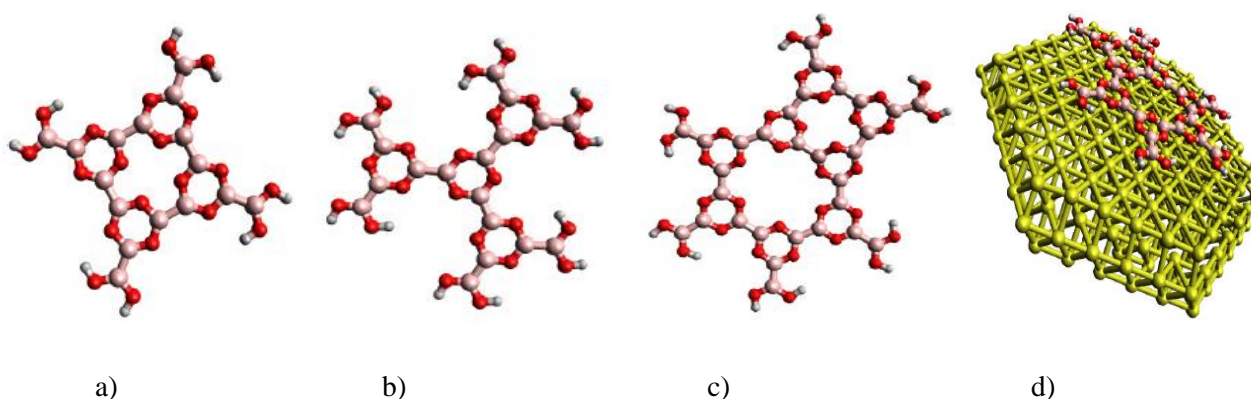
The B1s polarized experimental NEXAFS spectra of THDB on the Au(111) surface, measured in the work by M. Stredansky *et al.*<sup>14</sup>, display a pronounced dichroism, indicating a flat orientation of the 2D boroxine COF with respect to the Au(111) surface. In this thesis work, computational

simulations of B1s NEXAFS spectra of boroxine-based molecular models have been performed through DFT, in order to establish and quantify the relationship between the measured spectral features and the electronic structure information.



**Figure 5** – Experimental B1s (*s*- and *p*-pol) polarized NEXAFS spectrum of THDB on the Au(111) surface, measured by M. Stredansky et al.<sup>14</sup>.

Three molecular models (a, b, and c in Figure 6), made of boroxine rings, have been chosen to represent the 2D network on the Au(111) surface, with the aim of determining the best model to simulate the experimental spectral profiles. Model a) has been chosen because it represents a 4-fold motif which breaks the order of the regular structure in the STM image of Figure 3, while model b) is considered as the building block of the ideal periodic 2D network (lower panel of Figure 3). The third model (c) contain both the 4-fold and 6-fold motifs, and thus has been denoted as “4+6”; this model is also useful to describe the way in which the defects in the monolayer can affect the NEXAFS spectral profile of the regular structure. In analogy with the respective structural shape, model a) has been denoted as “square”, while model b) as “star”.



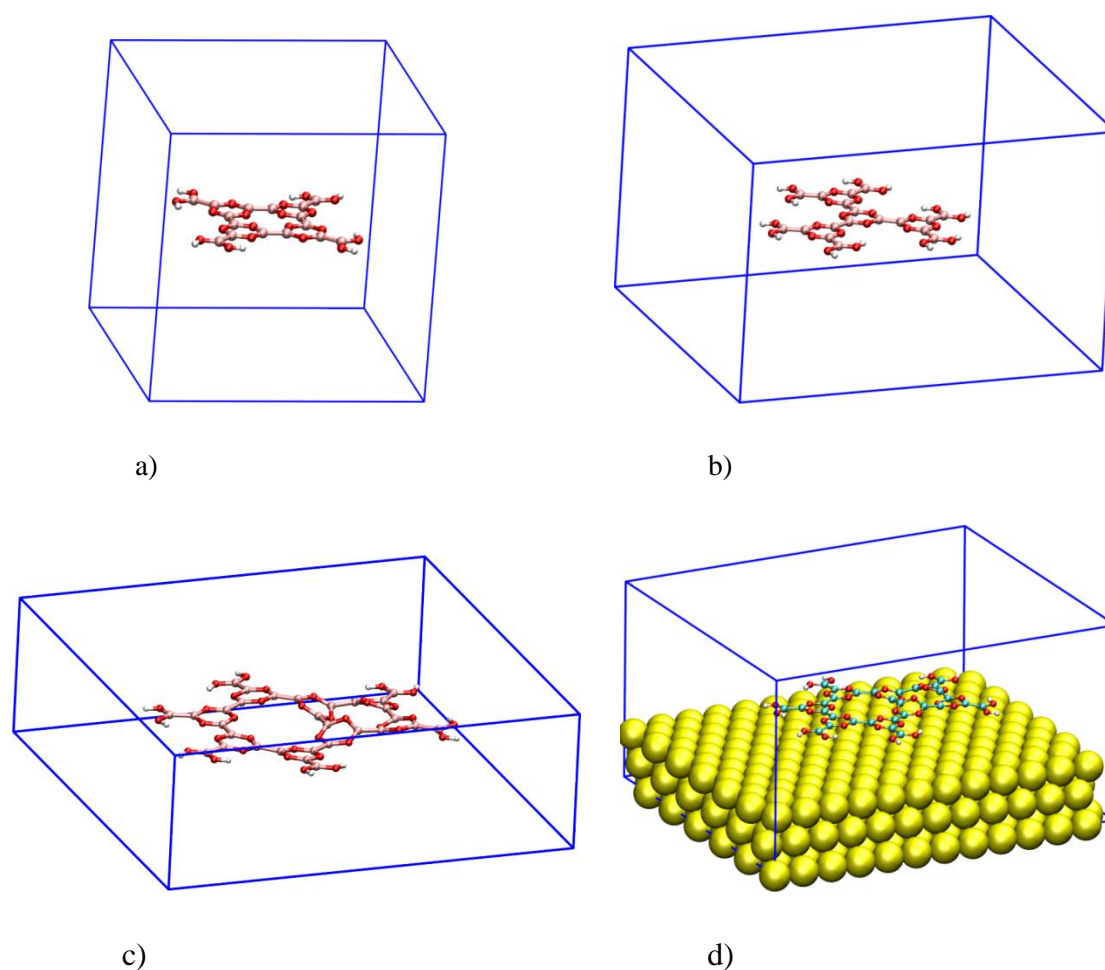
**Figure 6** – Ball-and-stick representation of the four models employed in B1s NEXAFS spectra calculations: a) “square”, b) “star”, c) “4+6” in gas-phase, d) “4+6” on Au(111). Link pink balls: B, Red balls: O, Grey balls: H, Yellow balls: Au.

All three models have been studied in the gas phase at the B1s NEXAFS *K-Edge*, *i.e.*, without considering any interaction with the Au(111) surface. In a second step, this interaction has been investigated by performing NEXAFS calculations on model c) adsorbed on a Au(111) slab. The cluster cut out from the optimized geometry of model c) adsorbed on Au(111) is reported in Figure 6d. We underline that the employment of a finite cluster for the adsorbed system allows the application of traditional DFT techniques to NEXAFS spectra simulations; moreover, the finite size cluster model allows to accurately represent events localized on the surface or within the bulk of solid systems. The nature of the interaction between the metallic surface and the boroxine-based model has been evaluated in terms of the differences between the spectral profile of the free (*i.e.*, gas phase) molecule and that on the molecule adsorbed on the surface: if the interaction between the adsorbate and the substrate is weak (physisorption), the relative spectral profile will be only slightly modified with respect to that of the gas phase molecule, thus preserving the electronic structure of the free species, while, if the overlap between the adsorbate and substrate molecular orbitals is more effective (chemisorption), a re-hybridization of the valence levels and consequent changes in the spectra will be observed, thus reflecting the influence of the surface in determining the electronic structure of the system.

## Computational Details

A two-step computational scheme has been adopted for the calculations of NEXAFS spectra of the systems considered. In the first step, a periodic slab methodology has been employed for the geometry optimizations of both periodic and isolated systems by means of the Quantum-Espresso code suite<sup>15</sup>, in collaboration with professor G. Balducci of the Department of Chemical and Pharmaceutical Sciences of the University of Trieste. Periodic calculations have been performed in the frame of Density Functional Theory (DFT)<sup>16</sup> with the Kohn-Sham (KS) orbitals expanded in a plane wave basis and the effects of atomic core regions described by ultrasoft pseudopotentials<sup>17</sup>. The (spin-unpolarized) generalized gradient approximation (GGA) in the PBE parameterization<sup>18</sup> has been adopted for modeling the exchange-correlation part of the energy functional. Van der Waals interactions have been taken into account by employing the vdW-DF functional<sup>19</sup>. The plane wave expansion of the crystalline orbitals has been truncated at a kinetic energy cutoff of 25 Ry, and a corresponding cutoff of 250 Ry has been used for the expansion of the augmentation charge needed by the ultrasoft pseudopotential method. The Brillouin zone integrations have been conducted at the  $\Gamma$  point, while the geometry optimizations have been performed by employing the BFGS (*Broyden-Fletcher-Goldfarb-Shanno*) algorithm<sup>20</sup>. For the geometry optimization of the

isolated molecular models, a single molecule has been (periodically) simulated in a unit supercell large enough to minimize any interactions between the molecule and its periodic images (greater than 10 Å between nearest atoms of any two contiguous periodic images). The Au(111) surface has been represented by a three-layer slab. The crystal slab was relaxed before docking onto it the (previously relaxed) isolated “4+6” molecule. For the relaxation of the “4+6” on Au(111) system, the molecule has been aligned parallel to the Au(111) surface, at about 4 Å above it; during the structural relaxation, the two bottom Au layers have been kept fixed, and those of the upper-most layer were allowed to vary together with the coordinates of the adsorbate. The supercells containing the investigated molecular models are shown in Figure 7.



**Figure 7** – Representation of the supercells used in the geometry optimization of the investigated molecular models: a) “square”, b) “star”, c) isolated “4+6”, d) “4+6” on the Au(111) slab.

In the second step, suitable finite clusters have been cut out from the periodic relaxed structures, and then used for the computation of NEXAFS spectra. Core excitation energies and oscillator strengths have been calculated by means of the molecular quantum chemistry Amsterdam Density Functional (ADF) program<sup>21</sup>. In order to get the electronic structure of the finite systems, scalar relativistic (SR) self-consistent field (SCF) KS calculations have been performed by using the

generalized gradient approximation (GGA) for the xc energy functional with the PW86x Perdew functional<sup>22</sup>. The Slater-Type Orbital (STO) atomic functions from the ADF database have been employed in the basis set. In particular, in NEXAFS spectra calculations of the gas phase systems, an even tempered quadruple  $\zeta$  with three polarization and three diffuse functions basis set (designed as ET-QZ3P-3DIFFUSE in the ADF database) has been used for the core-excited B atoms to describe the higher energy excitations that contribute to the near-edge structures, while a triple  $\zeta$  polarized (TZP) basis set has been employed for the remaining atoms. In particular, a frozen core (FC) TZP.1s basis set has been adopted for the B atoms not involved in the core excitation process, as well as for O atoms; the FC technique ensures the localization of the half core-hole. As concerns the surface cluster, the core-excited B atoms have been described by a ET-QZ3P-3DIFFUSE basis set, while the Au atoms directly interacting with the core-excited B atom have been treated with a zeroth-order regular approximation (ZORA) TZP basis set, freezing all core orbitals up to 4f (TZP.4f); for the remaining Au atoms of the cluster, a ZORA DZ.4f basis set has been used. In the case of the “4+6” system (both in gas phase and adsorbed on the Au(111) surface), the total number of basis functions has been lowered to obtain convergent results. In particular, a ET-QZP-3DIFFUSE basis set has been used the core-excited B atoms, and a TZP.1s basis set for core orbitals of the non-excited B atoms. A ZORA DZ.4f basis set has been used for all Au atoms of the surface; the employment of this basis set has been ultimately justified, since it has been found out that the interaction of the Au atoms with the adsorbed overlayer is weak.

The Transition Potential (TP) method<sup>23</sup> has been employed for the calculations of NEXAFS spectra; half an electron has been removed from the 1s core orbital of the excited B atom, leaving all virtual orbitals unoccupied and relaxing all orbitals until self-consistency. Excitation energies have been obtained as eigenvalues differences between the virtual orbital and the 1s core orbital:

$$\Delta E_{i \rightarrow f} = \epsilon_f^{TP} - \epsilon_i^{TP}, \quad (1)$$

while transition intensities are expressed in terms of oscillator strengths ( $f_{i \rightarrow f}$ ). For randomly oriented molecules in the gas phase:

$$f_{i \rightarrow f} = \frac{2}{3} n_i \Delta E_{i \rightarrow f} |\langle \varphi_f^{TP} | \boldsymbol{\mu} | \varphi_i^{TP} \rangle|^2 \quad (2)$$

involving dipole matrix elements between initial and final TP MOs, with  $n_i$  as occupation number of the core orbital in the ground state.

Expression (2) corresponds to the rotational average of that used in the case of fixed-in space molecules, according to which:

$$f_{i \rightarrow f} = 2n_i \Delta E_{i \rightarrow f} |\langle \varphi_f^{TP} | \hat{\mathbf{e}} \cdot \boldsymbol{\mu} | \varphi_i^{TP} \rangle|^2 \quad (3)$$

where  $\hat{\mathbf{e}}$  is the polarization vector of the incident radiation.

In the TP approach, the Ionization Potential (IP) corresponds to the negative of the TP eigenvalue related to the 1s core orbital,  $IP = -\epsilon_i^{TP}$ . More accurate values of the IPs have been obtained by performing  $\Delta$ KS ( $\Delta$ SCF Kohn-Sham) calculations allowing a full relaxation of the ionized core hole. The energy of the  $1s^{-1}$  ionic state has been extracted from a KS unrestricted calculation. The TP approach leads to a less attractive potential and the absolute transition energies are usually too large, therefore, the NEXAFS excitation energies (Eq. 1) have been adjusted by shifting them with respect to the  $\Delta$ IP value, corresponding to the energy difference  $[\epsilon_{1s}^{TP} - \Delta KS]$ .

All calculated spectral profiles have been convoluted with Gaussian functions of *Full Width at Half Maximum* (FWHM) equal to 0.3 eV, to achieve the best fit to the experimental measurements. In all NEXAFS spectra, oscillator strengths have been multiplied by a  $10^2$  factor. Additionally, for each system, oscillator strengths calculated for each nonequivalent B atom have been multiplied by the number of equivalent centers.

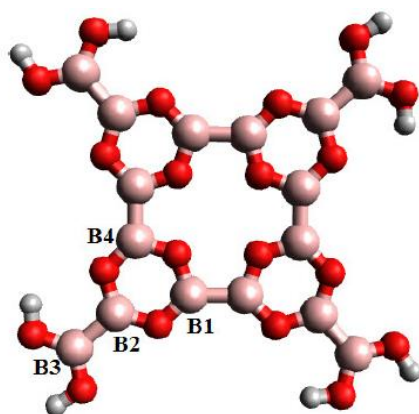
Finally, it is necessary to point out that the presented two-step computational methodology has been already effectively applied to organic molecules on Si(100)<sup>24,25</sup> and Au(111) surfaces<sup>26,27</sup>, proving its reliability for describing angle-resolved *K-edge* NEXAFS spectra of adsorbed molecules.

## Discussion and Results

In the following sections, the total calculated B1s NEXAFS spectra of the square, star and “4+6” gas-phase systems will be analyzed. Afterwards, a comparison between the total calculated B1s NEXAFS spectrum of the “4+6” gas-phase system and that relative to the adsorbate will be provided, in order to determine the extent of the Au-boroxine interaction. Finally, the total calculated B1s NEXAFS spectra of the square, star and “4+6” systems (the latter also adsorbed on the metal surface) will be compared with that relative to the THDB precursor (calculated in a previous thesis<sup>28</sup>, and with the experimental measurements performed on the adsorbed 2D COF<sup>14</sup>. This last comparison will be useful to evaluate the agreement between theory and experiment as well as the reliability of the employed molecular models.

### B1s NEXAFS spectrum of the square system

In Figure 8, a ball-and-stick representation of the “square” system is reported, together with the labelling of the nonequivalent B centers. In particular, four non-equivalent B core-hole sites have been identified; transition intensities have been therefore multiplied by 4 (the number of equivalent sites). The system has been assumed to belong to a *pseudo-C<sub>4h</sub>* point group, since its structure is not perfectly planar; indeed, in an ideal planar structure, a *C<sub>4</sub>* rotation axis and an inversion center *i* passing through the center of the inner ring of the molecule would be present.



**Figure 8** – Ball-and-stick representation of the “square” system with the labeling of nonequivalent B centers.



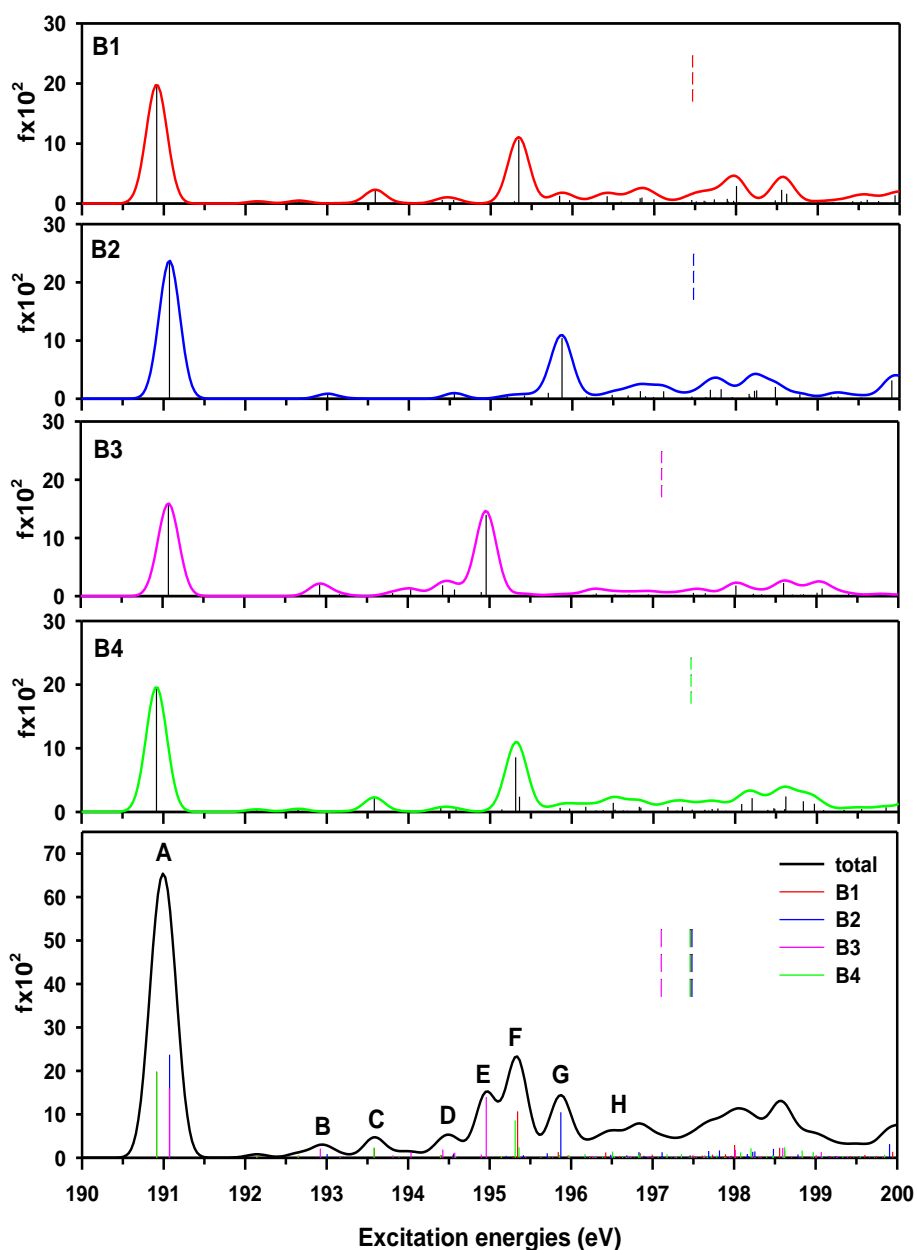
The  $\Delta$ KS IPs of each nonequivalent B site are collected in Table 1, where the B1s $\rightarrow$ LUMO excitation energies have been also included. The comparison between the  $\Delta$ KS IPs and B1s $\rightarrow$ LUMO excitation energies is useful to investigate the presence of *final state effects* (*i.e.*, orbital relaxation upon core-hole formation), which may lead to modifications in the LUMO excitation energy trend with respect to that displayed by the corresponding IPs.

**Table 1** – Comparison between B1s  $\Delta$ KS IPs and B1s $\rightarrow$ LUMO excitation energies of the “square” system. All values are expressed in eV.

Core-hole sites	$\Delta$ KS IPs	B1s $\rightarrow$ LUMO
B1	197.47	190.91
B2	197.48	191.07
B3	197.11	191.07
B4	197.46	190.91

From Table 1, one can evince that the IPs of the three nonequivalent B sites of each boroxine ring (*i.e.*, B1, B2, and B4) are all clustered in a narrow energy range, while the IP relative to B3 is the lowest one; this can be ascribed to *initial state effects* (*i.e.*, electronic density), due to the fact that B3, being part of the two terminal B(OH)<sub>2</sub> units, is less involved in the orbital delocalization which characterizes all four boroxine rings, and therefore results to be more shielded. Furthermore, the trend observed for the excitation energies associated to the B1s $\rightarrow$ LUMO transitions does not follow that of the corresponding IPs. Indeed, the excitation energies relative to the four LUMO transitions are in pairs degenerate, namely, B1-B4 and B2-B3. This can be explained as a final state relaxation effect of the LUMO upon the formation of the core-hole, which is quite different in the case of B2 and B3, but substantially similar for B1 and B4.

The total B1s NEXAFS spectrum of the “square” system is displayed in Figure 9, together with the partial contributions of each non-equivalent B<sub>*i*</sub> site (where the *i* index denotes the B atom carrying the core-hole).



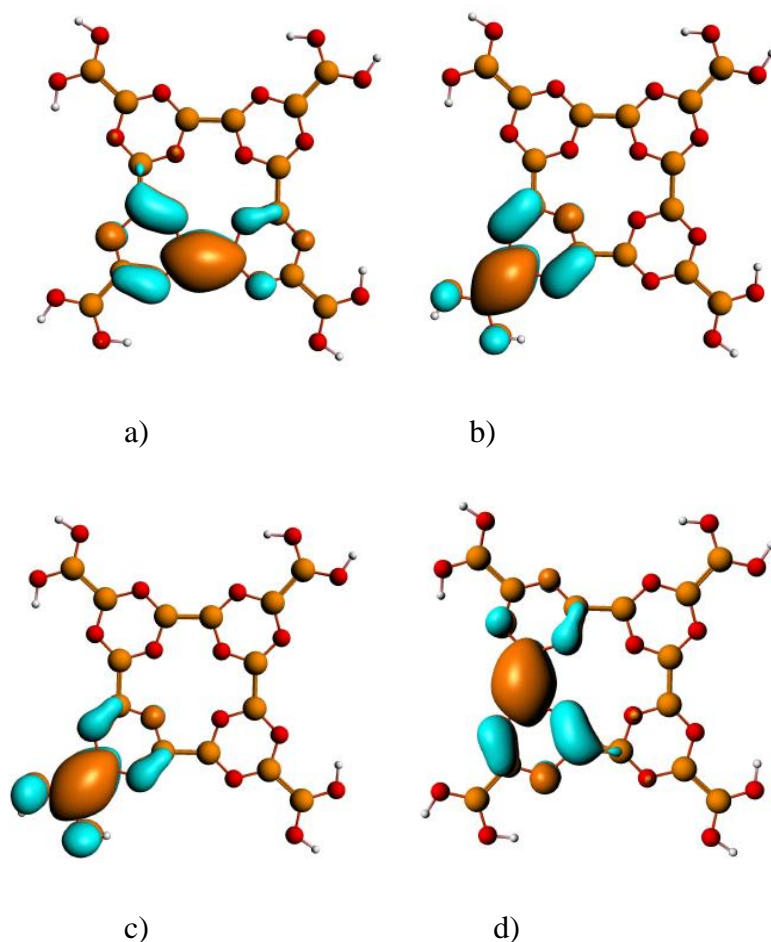
**Figure 9** – B1s NEXAFS spectra of the square system: the site-resolved B<sub>*i*</sub> contributions are reported in the first three panels, while the total calculated spectrum is displayed in the lower panel (black solid line) together with its deconvolution into the four B<sub>*i*</sub> partial contributions. The vertical colored dashed bars represent the  $\Delta$ KS IPs.

Each spectral feature has been analyzed through the aid of Table 2. Table 1 of the Appendix also reports the composition of the implicated MOs in terms of Mulliken population of AOs centered on each specific atom.

**Table 2** – Peak assignments of the total B1s NEXAFS spectrum of the “square” system. Only transitions with  $fx10^2 \geq 1.00$  are included.

Peaks	Core-hole site	E(eV)	$fx10^2$	Assignment
<b>A</b>	B1	190.91	19.8	LUMO/ $\pi$
	B4	190.91	19.6	
	B2	191.07	23.7	
	B3	191.07	16.0	
<b>B</b>	B3	192.92	2.11	LUMO+3/ $\pi^*$
<b>C</b>	B4	193.58	2.28	LUMO+5/ $\pi^*$
	B1	193.58	2.31	
<b>D</b>	B3	194.03	1.20	LUMO+7/mixed valence-Rydberg
	B3	194.42	1.89	LUMO+8/ $\pi^*$
	B3	194.58	1.17	LUMO+9/ mixed valence-Rydberg
<b>E</b>	B3	194.96	14.0	LUMO+11/mixed valence-Rydberg
<b>F</b>	B4	195.31	8.58	LUMO+10/mixed valence-Rydberg
	B1	195.34	10.7	LUMO+11/mixed valence-Rydberg
	B4	195.36	2.40	
<b>G</b>	B2	195.71	1.03	LUMO+12/ mixed valence-Rydberg
	B1	195.85	1.31	LUMO+14/mixed valence-Rydberg
	B2	195.87	10.5	LUMO+13/mixed valence-Rydberg
<b>H</b>	B1	196.42	1.23	LUMO+21/mixed valence-Rydberg
	B4	196.51	1.43	
	B2	196.83	1.33	LUMO+26/mixed valence-Rydberg
	B1	196.85	1.01	LUMO+27/mixed valence-Rydberg

At lower excitation energies, the total B1s NEXAFS spectrum is characterized by an intense band (A), which is contributed by the LUMO transitions starting from all four non-equivalent  $B_i$  core-hole sites; the LUMO orbital exhibits a  $\pi$   $B_i$ -B and B-O bonding character, and a  $\pi^*$   $B_i$ -O antibonding character, as one can observe from the relative plots (see Figure 10).



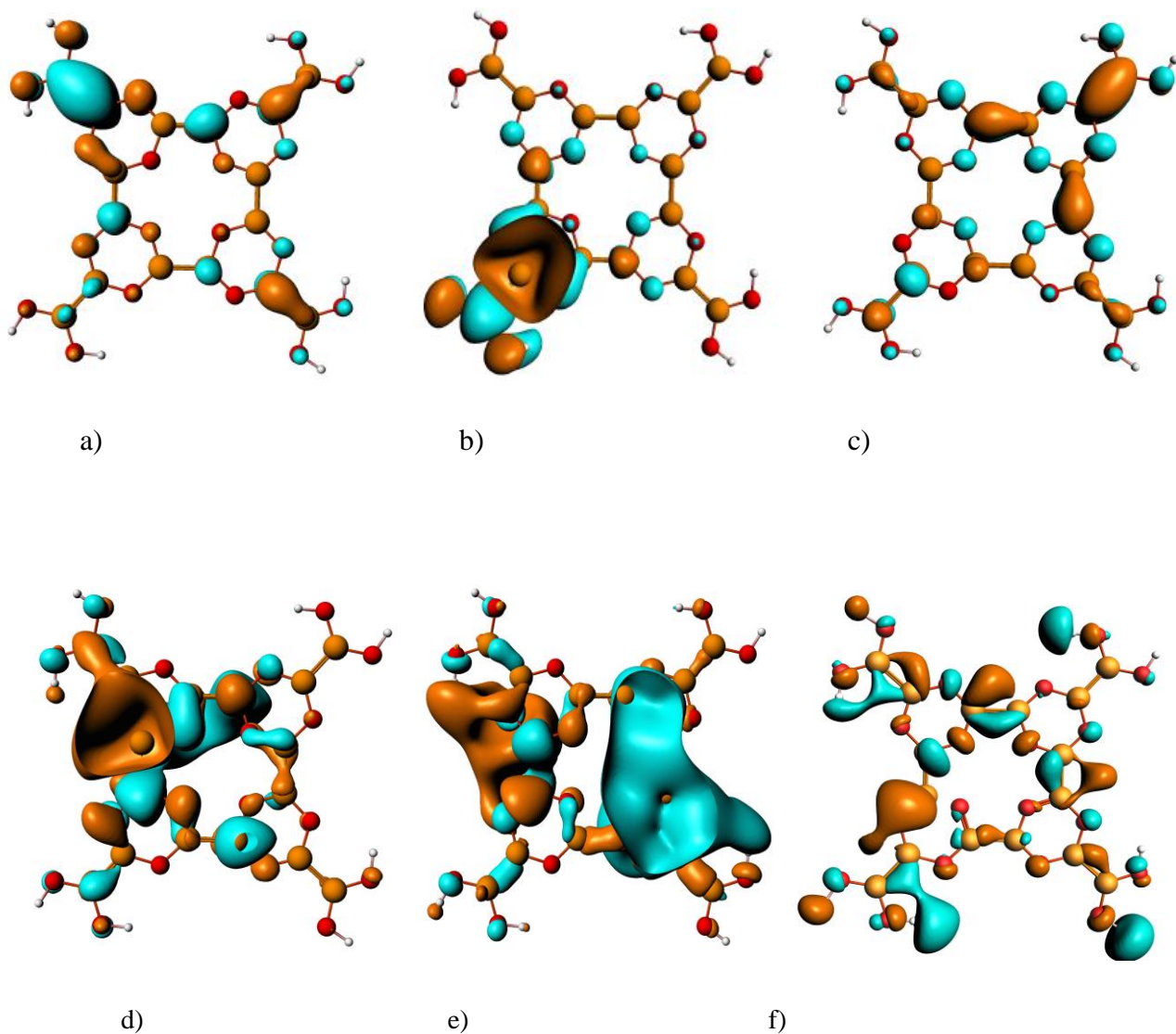
**Figure 10** – Plots of the LUMO orbitals relative to the B1 (a), B2 (b), B3 (c), B4 (d) transitions of the total B1s NEXAFS spectrum of the “square” system.

A detailed inspection of Figure 10 shows that the LUMOs relative to B1 and B4 transitions present essentially the same composition in terms of AOs, namely, strong  $2p_z$  contributions from the excited B atom and the neighboring B and O atoms, and smaller  $nd$  contributions from the excited B atom. The composition of the LUMO relative to the B2 transition shows a larger  $2p_z$  contribution from B2, with a consequent increase of the corresponding intensity. On the other hand, the intensity of the LUMO transition from the B3 site is lower, due to the smaller  $2p_z$  contribution from B3 (see Table 1 of the Appendix).

In the energy range between 192.5 and 194 eV, the total spectrum is characterized by two less intense structures (labelled B, and C) contributed by the transitions from all B<sub>i</sub> sites, with the exception of B2. The implicated transitions are toward MOs with  $\pi$  bonding B<sub>i</sub>-B character between B atoms not belonging to the ring where the core-hole is localized, and  $\pi^*$  B<sub>i</sub>-B antibonding character. In Figure 11a, the plot of a representative MO is shown.

Four more intense spectral features (labelled D, E, F, G) are visible in the energy region between 194 and 196 eV. Here, the total B1s NEXAFS spectrum reveals the sensitivity of NEXAFS spectroscopy to the chemical environment surrounding the core-hole, since the energy position of each spectral feature strongly depends on the localization of the core-hole. Indeed, looking at the partial contribution profiles, one can evince that the transitions from the B2 and B3 atoms differs significantly in energy, and this results in a broadening of the band in the total spectrum. On the other hand, the partial contribution profiles relative to B1 and B4 are very similar. All virtual MOs implicated in transitions giving rise to these four spectral features exhibit a mixed valence-Rydberg nature (see Figure 11b for the plot of a representative MO), with the exception of the B3 transition toward the LUMO+8, as one can observe from the relative plot displayed in Figure 11c. B4 contributes to band F with two transitions, namely, toward the LUMO+10 and LUMO+11 orbitals, of which, the latter exhibits a pronounced diffuse nature (see Figure 11d and 11e, respectively). The intensity of this band in the partial contribution profile of B4 is however comparable to that of the corresponding one in the partial contribution profile of B1; this confirms the formal equivalence between the two core-hole sites.

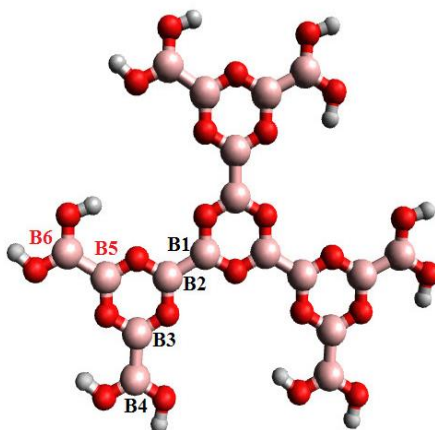
In the region close to the ionization threshold, a broader feature (H) is visible; it arises from a manifold of weak transitions from all non-equivalent core-hole sites toward virtual MOs with mixed valence-Rydberg nature, where the valence character is mainly given by  $\sigma^*(\text{O-H})$  contributions, as one can observe from the plot of MO relative to the B4 transition (see Figure 11f).



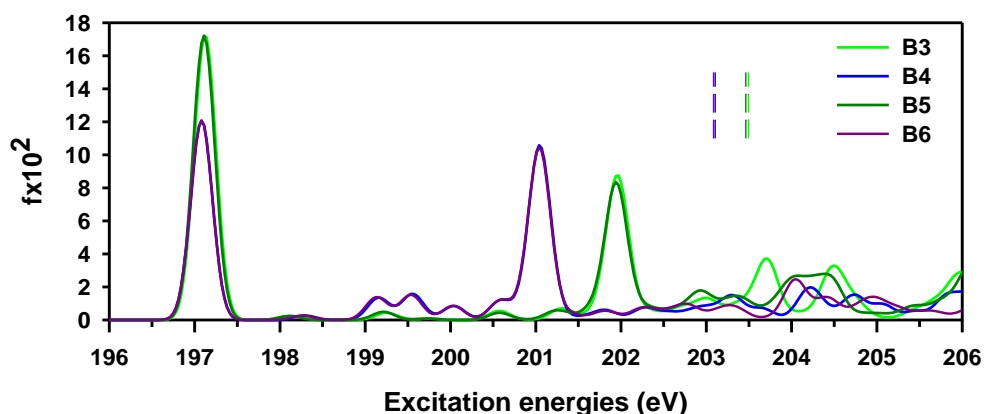
**Figure 11** – Plots of representative final MOs relative to the most intense transitions contributing to features B-G of the total B1s NEXAFS spectrum of the “square” system: a) B4→LUMO+5, b) B3→LUMO+11, c) B3→LUMO+8, d) B4→LUMO+10, e) B4→LUMO+11, f) B4→LUMO+21.

## B1s NEXAFS spectrum of the star system

In Figure 12, a ball-and-stick representation of the “star” system is shown, together with the labelling of the non-equivalent core-hole sites (black labels). Preliminary calculations on the B5 and B6 sites (red labels in Figure 12) have been performed to check their equivalence to the B3 and B4 sites, respectively; the related results are displayed in Figure 13.



**Figure 12** – Ball-and-stick representation of the star system with the labeling of nonequivalent B centers. The black labels denote the B atoms considered in NEXAFS spectra calculations, while the red labels indicate the B atoms used to check the equivalence between B centers.



**Figure 13** – Comparison between B1s NEXAFS spectral profiles of the B3, B4, B5, and B6 atoms of the “star” system. The vertical colored dashed bars represent the IPs calculated at the TP level.

From an inspection of Figure 13, it appears that the spectral profiles and  $IP^{TP}$  values of B3 and B4 are superimposable on those relative to B5 and B6, respectively, thus indicating their equivalence. Hence, only the four non-equivalent core-hole sites (B1, B2, B3 and B4) have been considered in the calculation of the B1s NEXAFS spectrum. Transition intensities have been therefore multiplied by 3 in the case of B1 and B2, and by 6 for B3 and B4. The system has been hence assumed to have a pseudo- $C_{3h}$  symmetry; indeed, a certain degree of bending is visible in its

structure. On the contrary, if the system was perfectly planar, a  $C_3$  rotation axis passing through the center of the inner ring of the molecule would be identified.

Before discussing the total calculated B1s NEXAFS spectrum, it is useful to discuss the trend of the  $\Delta$ KS IPs relative to the four non-equivalent B sites. IPs of each non-equivalent B site are reported in Table 3, together with the corresponding B1s $\rightarrow$ LUMO excitation energies.

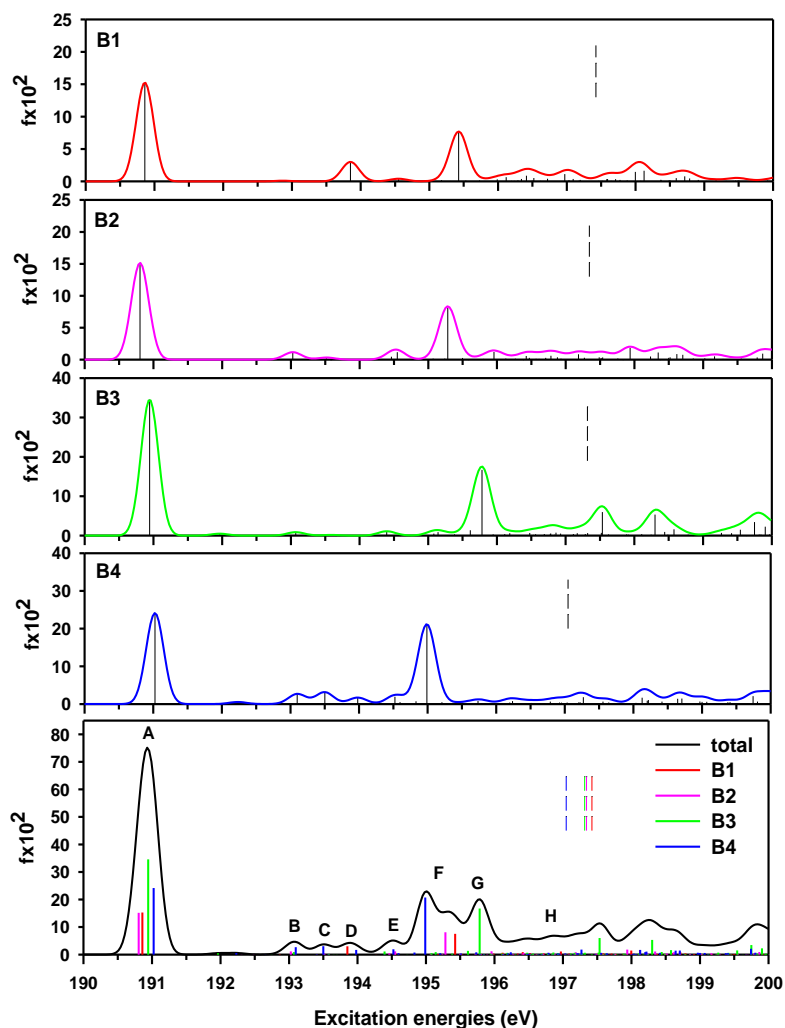
**Table 3** – Comparison between B1s  $\Delta$ KS IPs and B1s $\rightarrow$ LUMO excitation energies of the “star” system. All values are expressed in eV.

Core-hole sites	$\Delta$ KS IPs	B1s $\rightarrow$ LUMO
B1	197.42	190.86
B2	197.34	190.80
B3	197.32	190.94
B4	197.05	191.02

One can notice that the lowest IP value is predicted for the B4 site; this can be ascribed to *initial state effects*, which are strongly affected by the chemical environment of the core-hole. Indeed, B4 does not significantly participate to the orbital delocalization, which involves all three boroxine rings, and is therefore more shielded than the internal B atoms and its IP is the lowest. Looking at the excitation energies relative to the B1s $\rightarrow$ LUMO transitions, one can evince that the corresponding trend does not follow that of the IPs; therefore, *final state effects* are quite important, in particular ascribable to different relaxation among the B<sub>i</sub> core-hole sites of the virtual final states.



The total calculated B1s NEXAFS spectrum of the star system is shown in Figure 14, together with the partial contributions of each non-equivalent B site.



**Figure 14** – B1s NEXAFS spectra of the “star” system: the site-resolved B<sub>i</sub> contributions are reported in the first four panels, while the total calculated spectrum is displayed in the lower panel (black solid line) together with its deconvolution into the four B<sub>i</sub> partial contributions. The vertical colored dashed bars represent the  $\Delta$ KS IPs.

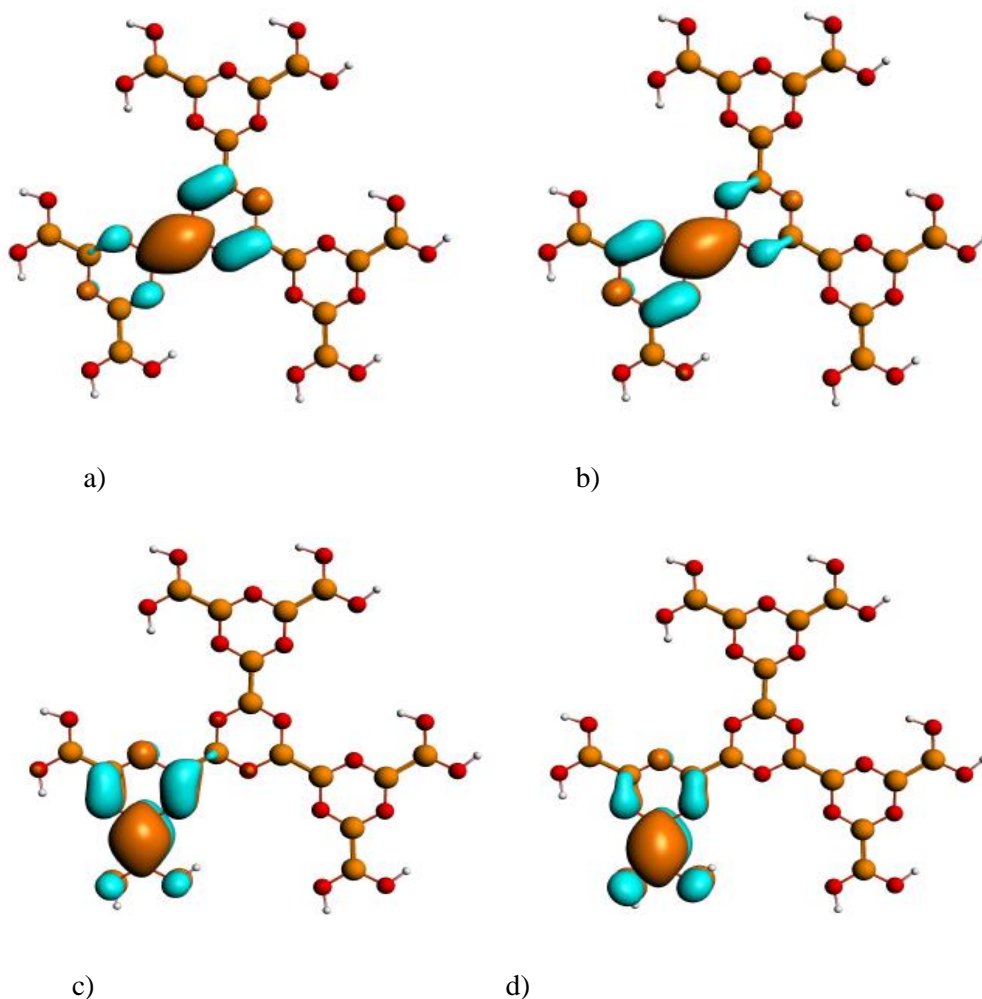
A first consideration on the total B1s NEXAFS spectrum regards the similarity to that relative to the “square system”, with an intense lower-energy peak (around 191 eV) dominating the spectrum and followed by weaker structures. Above 195 eV, a second intense feature is found; the energy position of the relative peaks (F-G in Figure 14) depends on the localization of the core-hole. The inspection of the B<sub>i</sub> partial spectra reveals that the terminal B4 atom gives a greater contribution to the total spectrum with respect to the other three core-hole sites.

The attribution of spectral features is guided by Table 4, which reports the most intense transitions contributing to the total spectrum and the assignment of the spectral features. In Table 2 of the Appendix, the composition of the involved MOs in terms of Mulliken population of AOs centered on each specific atom is also included.

**Table 4** – Peak assignments of the total B1s NEXAFS spectrum of the “star” system. Only transitions with  $f \times 10^2 \geq 1.00$  are included.

Peaks	Core-hole site	E(eV)	$f \times 10^2$	Assignment
<b>A</b>	B2	190.80	15.2	LUMO/ $\pi$
	B1	190.86	15.3	
	B3	190.94	34.6	
	B4	191.02	24.1	
<b>B</b>	B2	193.10	1.18	LUMO+3/ $\pi^*$
	B4	193.11	2.68	
<b>C</b>	B4	193.50	3.01	LUMO+5/ $\pi^*$
<b>D</b>	B1	193.85	3.01	LUMO+6/ $\pi^*$
<b>E</b>	B3	194.39	1.11	LUMO+7/mixed valence-Rydberg
	B4	194.52	1.92	LUMO+9/ $\pi^*$
	B2	194.54	1.18	
<b>F</b>	B4	194.99	20.7	LUMO+12/ $\pi^*$
	B2	195.28	8.09	
	B1	195.43	7.53	
<b>G</b>	B3	195.61	1.36	LUMO+13/mixed valence-Rydberg
	B3	195.68	16.7	LUMO+15/ $\pi^*$
<b>H</b>	B2	195.96	1.20	LUMO+18/mixed valence-Rydberg
	B1	196.97	1.13	LUMO+31/mixed valence-Rydberg

The first band (A) is assigned to the B1s $\rightarrow$ LUMO transitions from all four non-equivalent B<sub>i</sub> sites. The LUMO orbital displays a  $\pi$  B<sub>i</sub>-B bonding and  $\pi^*$  B<sub>i</sub>-O antibonding character, with strong 2p<sub>z</sub> AO contributions from the excited B atom and the neighbor O atoms; small *nd* AO contributions from the excited B atom are also involved (see Figure 15).

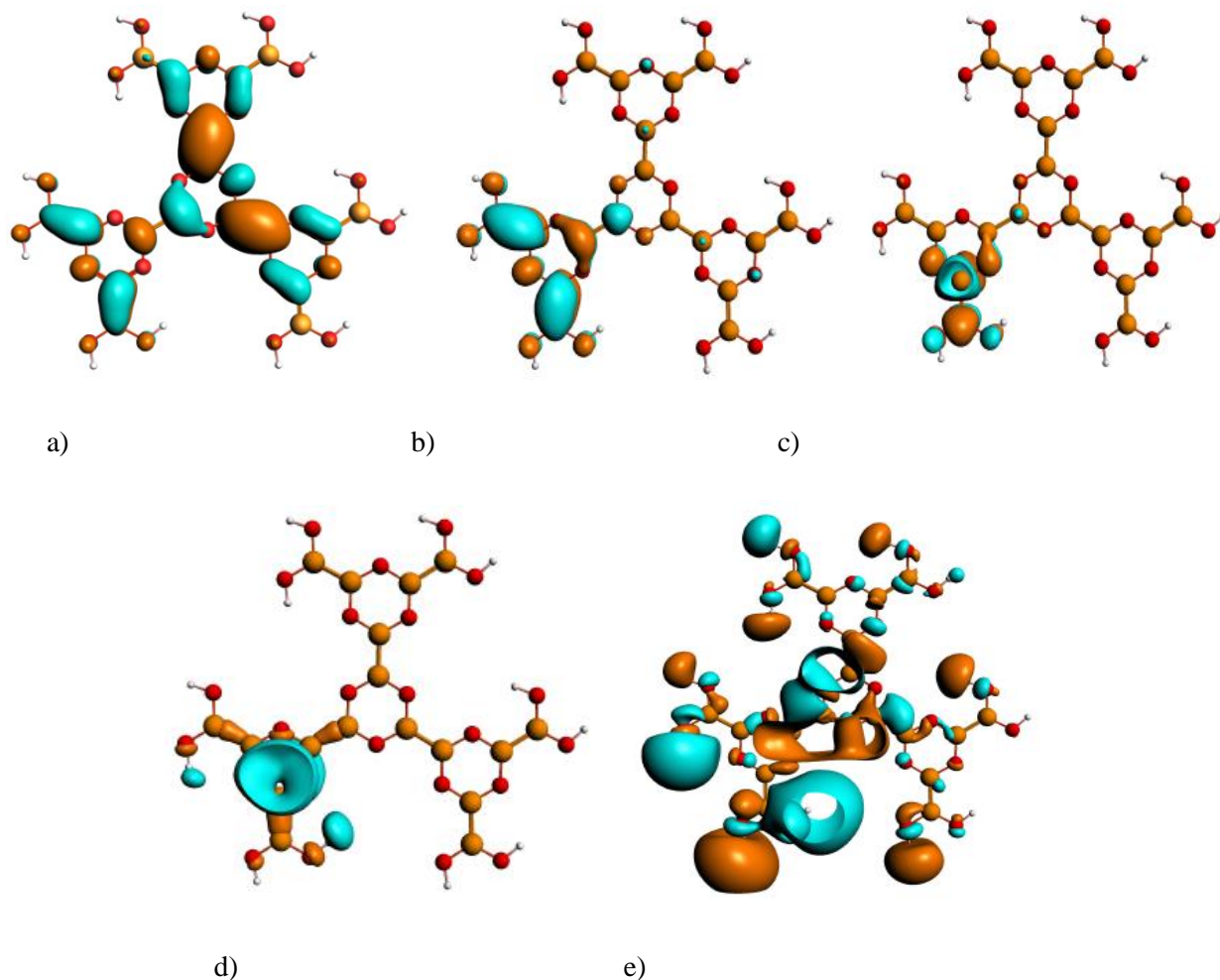


**Figure 15** – Plots of the LUMO orbitals relative to the B1 (a), B2 (b), B3 (c), B4 (d) transitions of the total B1s NEXAFS spectrum of the “star” system.

In the energy range between 193 and 194 eV, three weak features (labeled B, C, and D) are visible. Features B and C mainly derive from the B2 and B4 sites towards final MOs with  $\pi^*$  B<sub>i</sub>-B antibonding nature (see Figure 16a for the plot of a representative MO). Feature D is instead mainly contributed by a transition originating from the B1 toward the LUMO+6 orbital, which exhibits a  $\pi^*$  B<sub>i</sub>-B antibonding and  $\pi$  B3-B4 bonding character (see Figure 16b).

Three relatively intense spectral features (labelled E, F, and G) are found in the energy range between 194 and 196 eV. These three features are mainly contributed by transitions toward final MOs of  $\pi^*$  B<sub>i</sub>-B antibonding nature, with small diffuse *nd, np* AO contributions from the excited B atom, as one can observe from the plot of the LUMO+12 relative to the B4 transition (see Figure 16c). Transitions toward mixed-valence Rydberg MOs also occur in this energy range; as representative examples, the plot of the LUMO+7 orbital relative to the B3 transition is shown in Figure 16d: it presents a strong diffuse nature with small 2p<sub>y</sub> AO contributions from the B atoms surrounding the excited one.

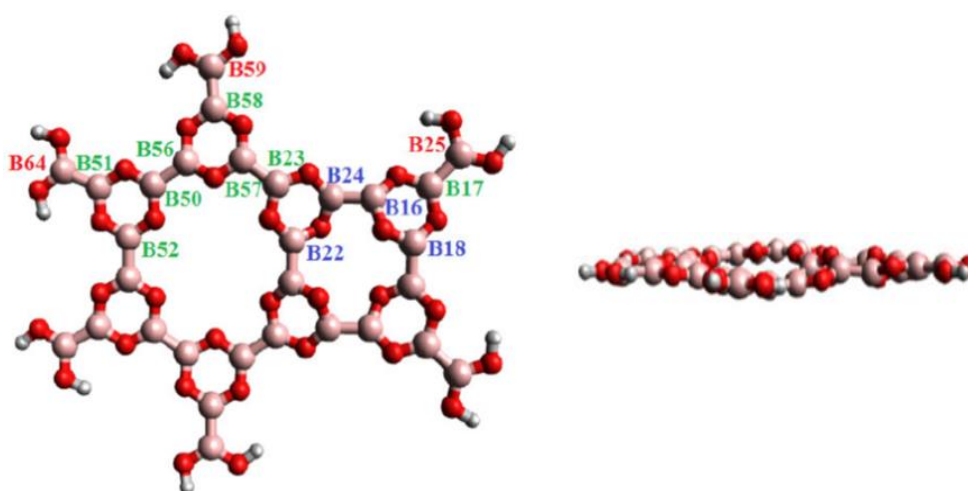
Finally, a broader feature (H) is visible in the region close to the ionization threshold: it arises from a manifold of weak transitions from all four non-equivalent core-hole sites; among these, the most intense transitions are relative to B1 and B2, and are toward final MOs of mixed valence-Rydberg character. The valence nature of these MOs mainly involves  $\sigma^*(\text{O-H})$  contributions, as one can observe from the plot of the MO relative to the transition from B1, reported in Figure 16e.



**Figure 16** – Plots of representative final MOs relative to the most intense transitions contributing to features B-H of the total B1s NEXAFS spectrum of the “star” system: a) B2→LUMO+3, b) B1→LUMO+6, c) B4→LUMO+12, d) B3→LUMO+7, e) B1→LUMO+31.

## B1s NEXAFS spectrum of the “4+6” gas-phase system

A ball-and-stick representation of the “4+6” gas-phase system is shown in Figure 17 (left side), together with the labeling of the non-equivalent B<sub>i</sub> sites. In particular, fifteen non-equivalent B centers have been identified, and further subdivided into three groups, including *terminal*, *distorted* and *planar* atoms, according to the corresponding position and arrangement in the molecule. The system has been assumed to have a *pseudo-C*<sub>2</sub> symmetry, since its structure slightly diverges from planarity (see Figure 17, right side); indeed, in an ideal planar structure, a C<sub>2</sub> rotation axis passing through the center of the square and hexagon would be visible. For each non-equivalent B site, transition intensities have been multiplied by 2 (the number of equivalent centers).



**Figure 17** – Left side: ball-and-stick representation of gas-phase “4+6” (top view) with the labeling of B non-equivalent centers (red: *terminal* atoms, blue: *distorted* atoms, green: *planar* atoms). Right side: Side view of gas-phase “4+6” showing the degree of bending.

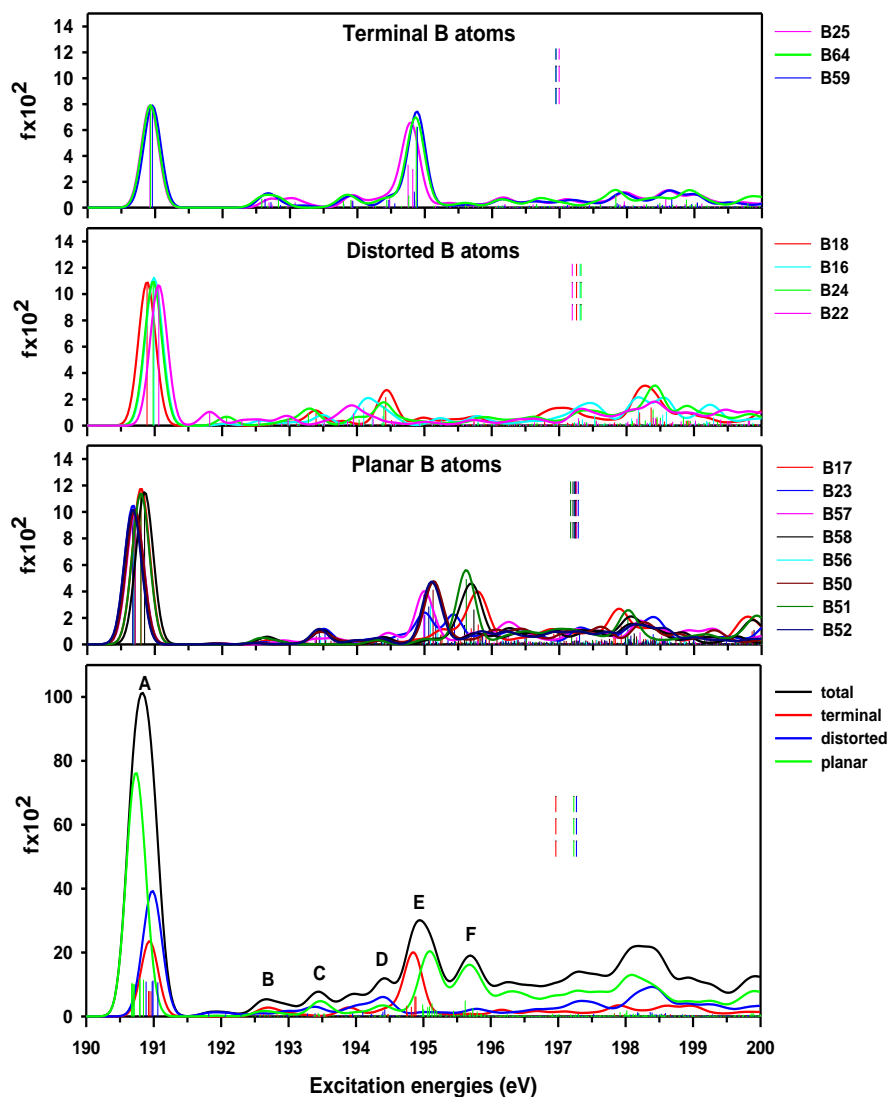
We first rationalize the IP trend displayed by the three groups of non-equivalent B sites.  $\Delta$ KS IPs of each non-equivalent B site are reported in Table 5.

**Table 5** –  $\Delta$ KS IPs (in eV) of each non-equivalent B site of the “4+6” gas-phase system.

<b>Terminal</b>		<b>Distorted</b>		<b>Planar</b>	
B64	196.944	B22	197.194	B51	197.167
B59	196.957	B18	197.257	B58	197.204
B25	196.993	B16	197.310	B56	197.222
		B24	197.324	B52	197.229
				B50	197.242
				B57	197.250
				B17	197.251
				B23	197.282

The IPs of the *terminal* B atoms are the lowest ones, as observed for the “square” and the “star” systems; this can be ascribed to *initial state* effects, as these sites does not belong to the boroxine ring, and therefore are less involved in the orbital delocalization which characterizes all rings of the molecule. On the other hand, the IP values related to both *planar* and *distorted* B atoms do not follow a clear trend. Another point to highlight is that the energy order of the IPs does not reflect that of the excitation energies of the corresponding LUMO transitions (reported in Table 6); this points toward the presence of *final state* effects. In particular, it is possible to observe that the lowest excitation energies of the LUMO transitions are observed for the *planar* atoms, whereas the lowest IP values are found for the *terminal* atoms, whose excitation energies relative to the LUMO transitions are comprised between those of the *planar* and *distorted* atoms (compare Table 5 and Table 6).

The total B1s NEXAFS spectrum of the “4+6” gas-phase system is displayed in Figure 18, together with the partial contributions of each nonequivalent B site.



**Figure 18** – B1s NEXAFS spectra of gas-phase “4+6”: the site-resolved  $B_i$  contributions are reported in the first three panels, while the total calculated spectrum is displayed in the lower panel (black solid line) together with its deconvolution into the three  $B_i$  partial contributions. The vertical colored dashed bars represent the  $\Delta$ KS IPs.

A first consideration concerning the total B1s NEXAFS spectrum of the “4+6” gas-phase system regards the similarity to the spectral profiles of the “square” and the “star” systems: a first intense peak around 191 eV followed by weak transitions in the 192-194 eV energy range, and a more intense peaked features around 195 eV. Another general observation is related to the different shape of the partial contributions relative to the *distorted* atoms with respect to that displayed by the other two groups of core-hole sites; this can be rationalized by considering that the *distorted* atoms feel the molecular strain of the whole “4+6” system, which determines the divergence from planarity. As a result, a depression of the intensity is observed for transitions starting from these

sites in the energy region around 195 eV, along with the different nature of the implicated final MOs. Each spectral feature of the total spectrum has been assigned through the aid of Table 6. In Table 3 of the Appendix, the composition of the involved MOs in terms of Mulliken population of AOs centered on each specific atom is also reported.

**Table 6** – Peak assignments of the total B1s NEXAFS spectrum of the “4+6” gas-phase system. Only transitions with  $fx10^2 \geq 0.40$  are included.

Peaks	Core-hole site	E(eV)	fx10 <sup>2</sup>	Assignment
<b>A</b>	B52	190.67	10.2	LUMO/ $\pi$
	B23	190.68	10.5	
	B56	190.69	10.1	
	B50	190.71	10.2	
	B57	190.71	10.0	
	B51	190.79	11.4	
	B17	190.80	11.8	
	B58	190.85	11.5	
	B18	190.89	10.9	
	B25	190.92	7.91	
	B64	190.93	7.93	
	B59	190.96	7.90	
	B24	190.98	11.0	
	B16	190.99	11.2	
B22	191.06	11.7		
<b>B</b>	B59	192.58	0.67	LUMO+2/ $\pi^*$
	B25	192.59	0.41	
	B64	192.63	0.68	
<b>C</b>	B24	193.28	0.95	LUMO+7/ $\pi^*$
	B18	193.39	0.96	LUMO+8/ $\pi^*$
	B23	193.52	0.98	LUMO+9/ $\pi^*$
<b>D</b>	B22	193.95	1.93	LUMO+12/valence
	B16	194.10	1.37	LUMO+11/valence
	B16	194.23	0.79	LUMO+12/valence
	B24	194.40	1.51	LUMO+14/mixed valence-Rydberg
	B18	194.43	2.18	LUMO+13/mixed valence-Rydberg
<b>E</b>	B25	194.75	3.31	LUMO+16/valence
	B25	194.82	2.98	LUMO+17/mixed valence-Rydberg
	B59	194.85	1.24	
	B64	194.87	6.26	LUMO+18/mixed valence-Rydberg
	B59	194.89	6.25	



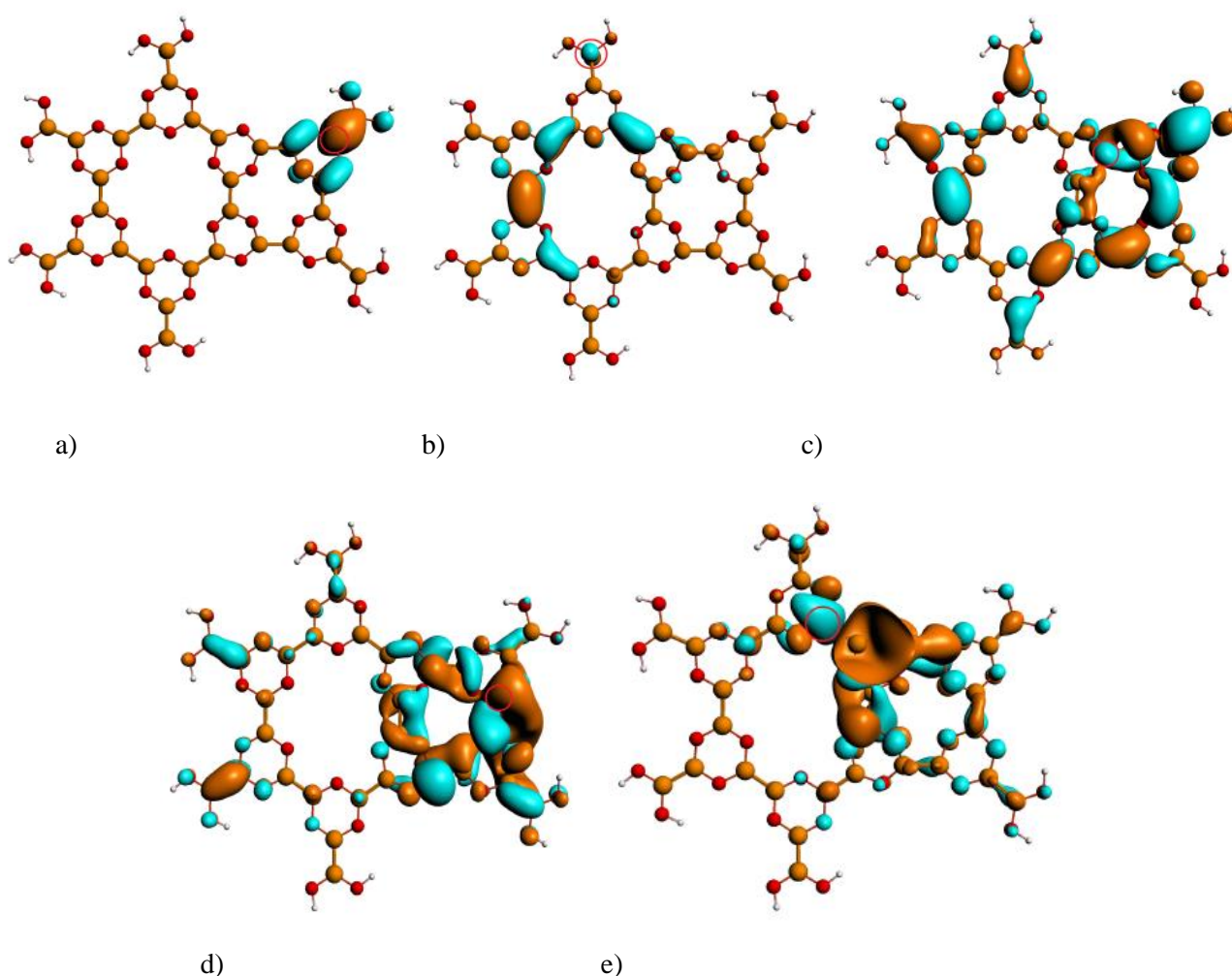
	B57	194.99	3.78	LUMO+16/mixed valence-Rydberg
	B23	195.00	2.24	LUMO+17/mixed valence-Rydberg
	B52	195.06	2.87	
	B56	195.09	3.12	
	B50	195.13	4.13	
	B52	195.17	2.19	LUMO+18/mixed valence-Rydberg
	B56	195.20	1.85	
<b>F</b>	B23	195.45	1.81	LUMO+21/mixed valence-Rydberg
	B58	195.60	1.47	
	B51	195.62	4.94	
	B17	195.69	1.22	LUMO+22/mixed valence-Rydberg
	B58	195.73	2.64	LUMO+23/mixed valence-Rydberg
	B17	195.80	1.51	LUMO+24/mixed valence-Rydberg

The first intense absorption band (A) arises from the LUMO transitions starting from all fifteen non-equivalent sites; in particular, transitions from the *planar* atoms occur at lower excitation energies, whereas transitions related to the other two groups of core-hole sites contribute to the higher-energy portion of the band. The plot of a LUMO related to a transition from a planar atom (B17) is displayed in Figure 19a; this MO exhibits a  $\pi$  B<sub>i</sub>-B bonding and  $\pi^*$  B<sub>i</sub>-B antibonding character with strong 2p<sub>z</sub> contributions from the B atom carrying the core-hole, and lower 2p<sub>z</sub> contributions from the neighboring B and O atoms. The latter contribute to a larger extent to the LUMOs relative to *terminal* B atoms, while smaller 2p<sub>z</sub> contributions from the core-excited B atom are detected (see Table 3 of the Appendix); this explains the lower intensity of the implicated transitions with respect to those from the other two sets of core-hole sites.

The following two bands (B and C) are much less intense than the first one due to the smaller 2p<sub>z</sub> AO contributions from the B atoms carrying the core-hole to the implicated final MOs (see Table 3 of the Appendix). In particular, band B is contributed by transitions originating from *terminal* B atoms, while band C mainly arises from core-excitations of the *planar* and *distorted* B atoms. Representative examples of the involved MOs are reported in panels b), and c) of Figure 19.

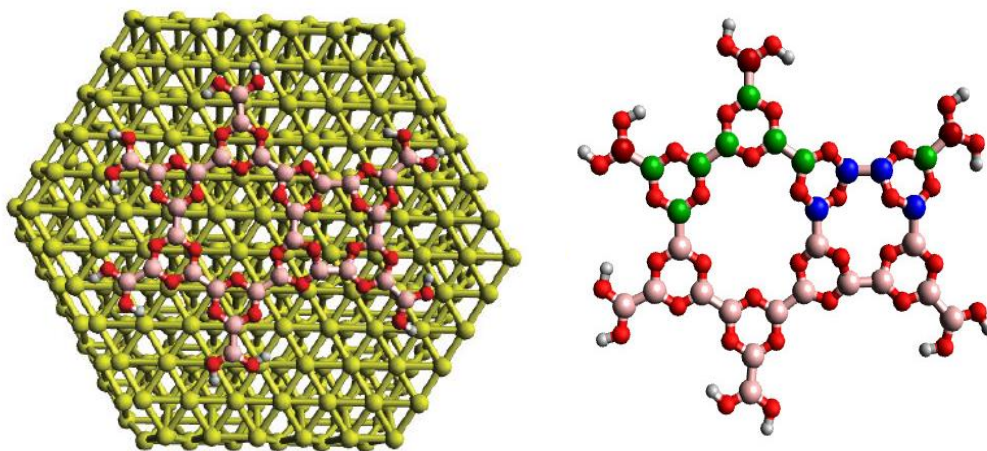
The most interesting features (labelled D, E, and F) fall in the energy range between 194 and 196 eV; the energy position of these features is strongly affected by the localization of the core-hole. As revealed by the B<sub>i</sub> partial spectral profiles (see Figure 18, upper panels), band D is assigned to transitions from the *distorted* B atoms; it is worth noting that this spectral feature was not visible in the total B1s NEXAFS spectra of the “square” and “star” systems, in line with the lack of this

typology of B sites in those systems. The implicated transitions are toward MOs displaying purely valence or mixed valence-Rydberg character. In Figure 19d, the plot of a representative MO, relative to B18, is shown; as one can observe, it is mainly delocalized over the square. Band (E), centered at around 195 eV, is ascribed to transitions starting from the *terminal* and *planar* B atoms, contributing to the lower and higher-energy portion of the band, respectively. The involved MOs exhibit a mixed valence-Rydberg character, with  $\pi^*$  B<sub>1</sub>-B antibonding contributions and diffuse components, as it is possible to observe from the plot of the LUMO+16 relative to the B57 transition (see Figure 19e). The last intense structure (F) in the energy region close to the ionization threshold is attributed to core-excitations of the *planar* B atoms toward mixed valence-Rydberg MOs.



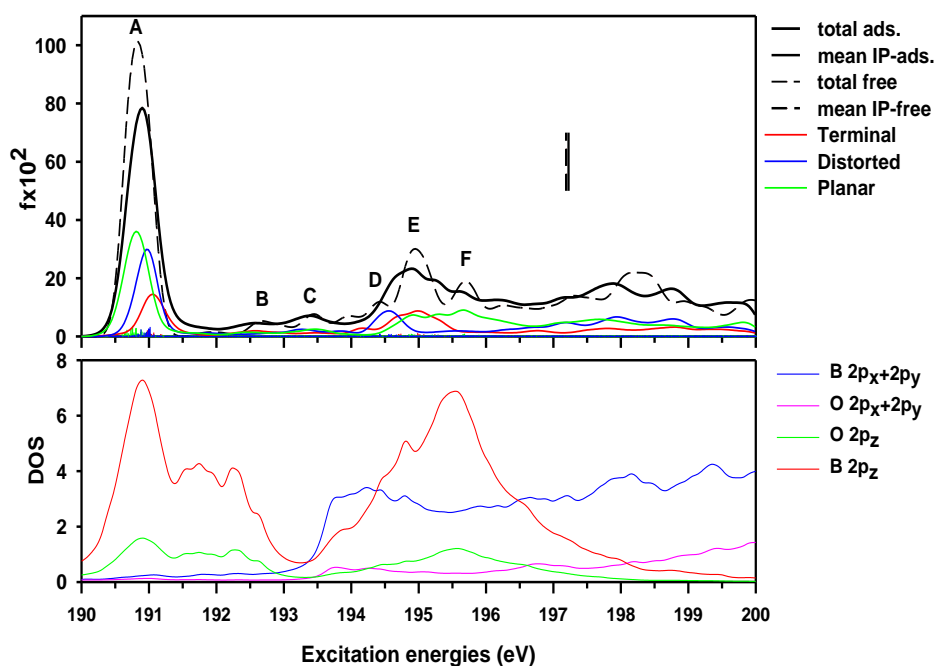
**Figure 19** – Plots of representative final MOs implicated in the most intense transitions contributing of the total B1s NEXAFS spectrum of the gas-phase “4+6” system: a) B17→LUMO, b) B59→LUMO+2, c) B24→LUMO+7, d) B18→LUMO+13, e) B57→LUMO+16. The orange circles identify the B atom which is excited in each transition.

In Figure 20, a ball-and-stick representation of the “4+6” system adsorbed on the Au(111) surface is displayed.



**Figure 20** – Left side: ball-and-stick model of “4+6” adsorbed on Au(111) (top-view). Right side: ball-and-stick representation of free “4+6” showing the nonequivalent B sites with different colors (dark red: terminal B atoms, blue: distorted B atoms, green: planar B atoms).

A comparison between the total B1s NEXAFS spectrum of the “4+6” gas-phase system and that relative to the same system adsorbed on Au(111) is visible in Figure 21. As in the case of the gas-phase system, fifteen non-equivalent B sites have been considered in the computation of NEXAFS spectra. In the lower panel of Figure 21, the partial density of the virtual states (PDOS) calculated with the ground state configuration for the 2p contributions of the B and O atoms of the adsorbed system is reported. This is useful for the analysis of the virtual MOs which represents a difficult task due to the large extension of the system and high density of excited states. The computed PDOS profiles can be compared to the corresponding NEXAFS spectrum, since, within a single particle approximation, one can assume that the core excitation spectrum is associated with the partial density of the virtual states of the atom carrying the core-hole which is dipole-allowed. PDOS profiles have been shifted to higher excitation energies to get an easier comparison with the NEXAFS spectrum; in particular, the first intense band of the  $2p_z$  PDOS profile relative to the B atoms has been aligned with the first band of the corresponding NEXAFS spectrum, due to the significant domain of the B  $2p_z$  contributions in transitions occurring in this energy range; the same procedure has been adopted for the remaining PDOS profiles.



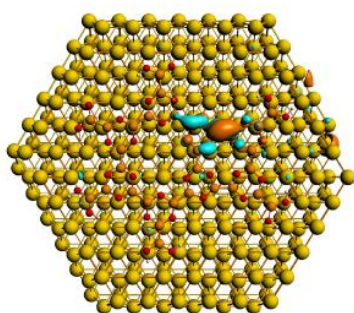
**Figure 21** – Upper panel: Comparison between the total B1s NEXAFS spectrum of “4+6” free (black solid line) and adsorbed on Au(111) (black dashed line). Partial  $B_i$  contributions of the three groups of nonequivalent B atoms are also shown. Vertical dashed bars: mean  $\Delta$ KS IPs (equal to 197.19 eV and 197.22 eV for “4+6” free and adsorbed on Au(111), respectively). Lower panel: PDOS profiles relative to the  $2p_x$ - $2p_y$  and  $2p_z$  partial contribution of the B and O atoms of “4+6” adsorbed on Au(111). PDOS profiles have been convoluted with Lorentzian function with FWHM equal to 0.12 eV.

From a detailed analysis of the PDOS profiles (lower panel of Figure 21), one can observe that the O  $2p_x$ ,  $2p_y$  and  $2p_z$  atomic contributions follow the same trend of the corresponding B atomic orbitals, although displaying a lower intensity due to the smaller participation to the implicated final MOs. PDOS relative to the Au atoms have been not reported, since the analysis of the empty states in the energy range between 190 and 200 eV has shown that the most significant AO contributions are those forming the  $sp$  band, but with a weak mixing with MOs contributed by the AOs of the molecule.

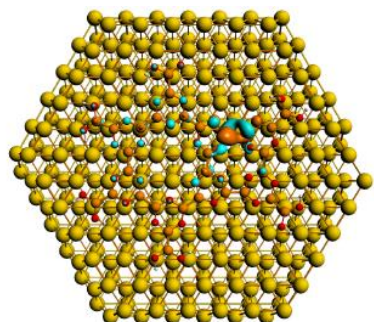
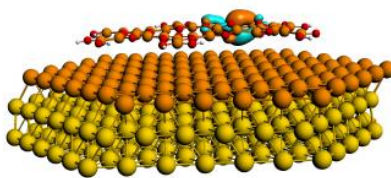
As concerns the comparison of the total calculated B1s NEXAFS spectrum of the gas-phase system with that relative to the system adsorbed on Au(111) (upper panel of Figure 21), it appears that the two spectral profiles do not display significant variations in the intensity distribution. Moreover, it is possible to notice that the intensity of the first band decreases upon adsorption, and shows a more asymmetric shape, which is related to the redistribution of the intensity into several less intense transitions, involving virtual MOs presenting a certain contribution from Au AOs.

The virtual MO implicated in the most intense transition of the adsorbed system, giving rise to band A, exhibits a  $\pi$  nature, and is entirely localized on the adsorbate (see Figure 22a). Bands B and C essentially maintain the same energy position and intensity in both NEXAFS spectra, whereas the

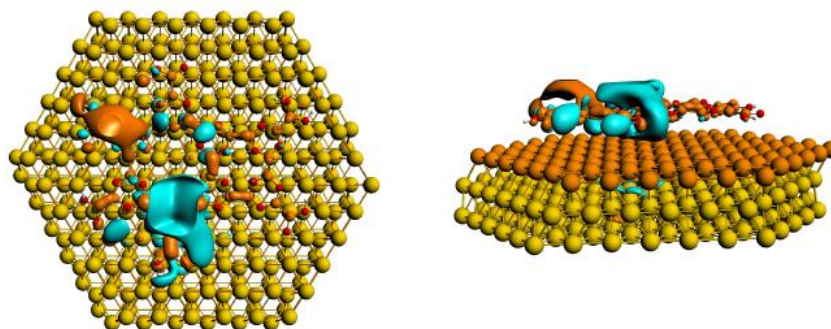
three higher-energy bands (labelled D, E, and F) of the spectrum of the free molecule are encompassed in a lower-intensity single smooth band in the spectrum relative to the adsorbate. It is also interesting to point out that, at the lower energy side of this band, the B  $2p_x$  and  $2p_y$  AO contributions increase; this is consistent with the significant participation of the *distorted* B atoms to the final MOs implicated in the corresponding transitions. These MOs are rehybridized and exhibit mixed valence-Rydberg character, as in the case of the gas-phase molecule; this is clearly visible in the plot of the final MO associated with the intense transition at around 194.6 eV (see Figure 22b). The implicated MO is delocalized over the entire molecule, and does not show any contribution from Au AOs. The higher-energy portion of the band (194.5 eV-195.5 eV) can be instead ascribed to transitions starting from the *planar* and *terminal* B atoms, while the remaining portion (195.5 eV-196 eV) is attributed exclusively to the *planar* B atoms, in line with bands E and F of the spectrum of the gas-phase molecule. In this energy range, the B  $2p_x$  and  $2p_y$  contribution is essentially the same, whereas the B  $2p_z$  one significantly increases up to 195.5 eV and then progressively decreases. In Figure 22c, the plot of a MO involved in an intense transition occurring in this energy range is visible. The latter mainly exhibits  $2p_z$  and  $np$  AO contributions from the B atoms; Au AOs are instead not involved.



a)



b)

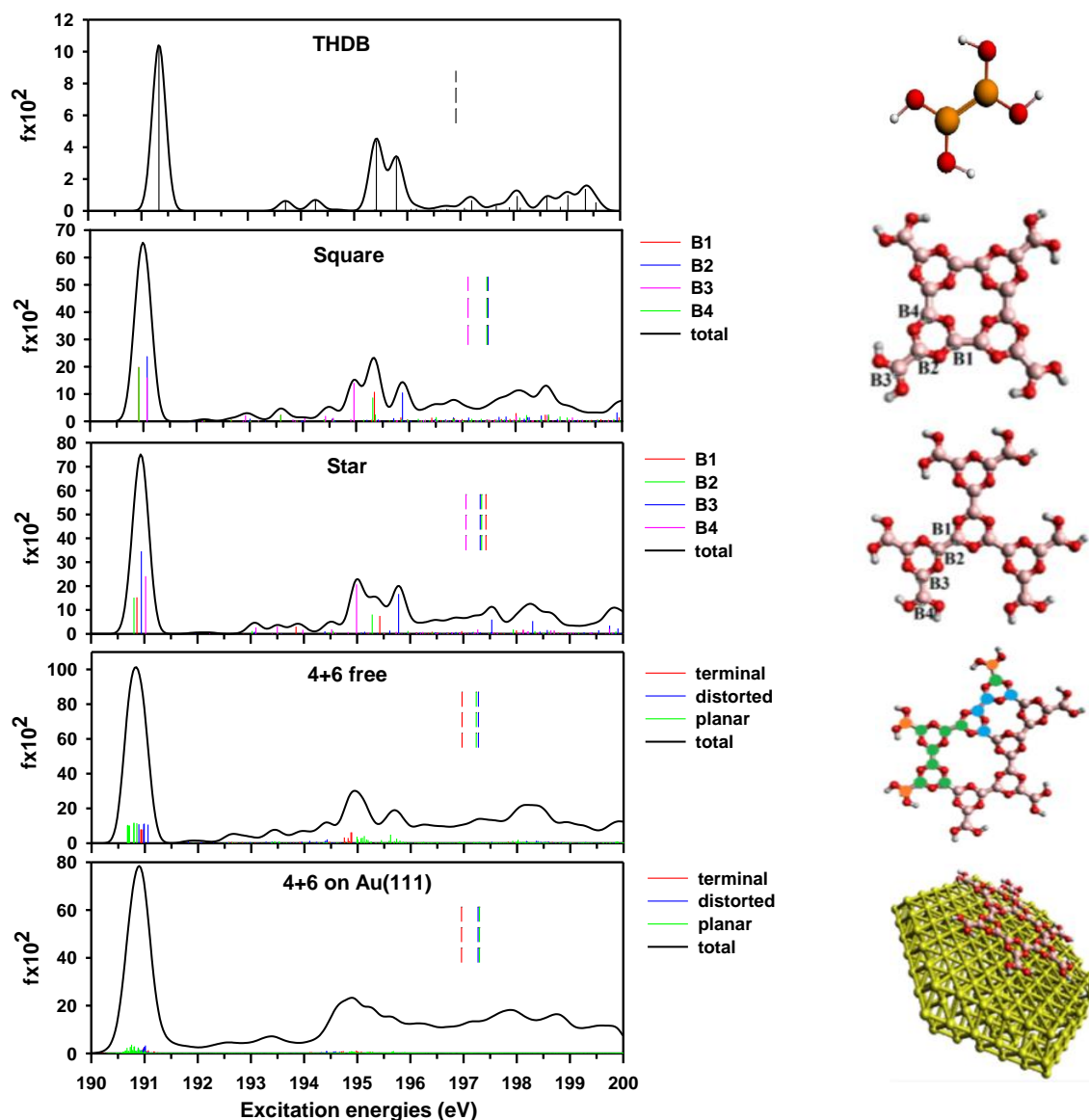


c)

**Figure 22** – Plots of representative virtual MOs (left: top view, right: side view) implicated in the most intense transitions relative to core excitations in the “4+6” system adsorbed on Au(111): a) B24→LUMO+29, b) B16→LUMO+241, c) B50→LUMO+283.

Finally, since the spectral profile of the gas-phase molecule is similar to that relative to the adsorbate, it is possible to conclude that the Au-boroxine interaction is weak (*i.e.*, physisorption); this has been also confirmed by the strong contribution of B,O AOs to the final MOs involved in the most intense transitions giving rise to the spectrum of the adsorbate (see Figure 21).

In Figure 23, the total calculated B1s NEXAFS spectra of the three investigated free molecules and of the “4+6” system on the Au(111) surface are compared; also the spectrum of the THDB precursor is included, in order to study the evolution of spectral features by increasing the molecular complexity, *i.e.*, the presence of defects and the Au-B interaction.



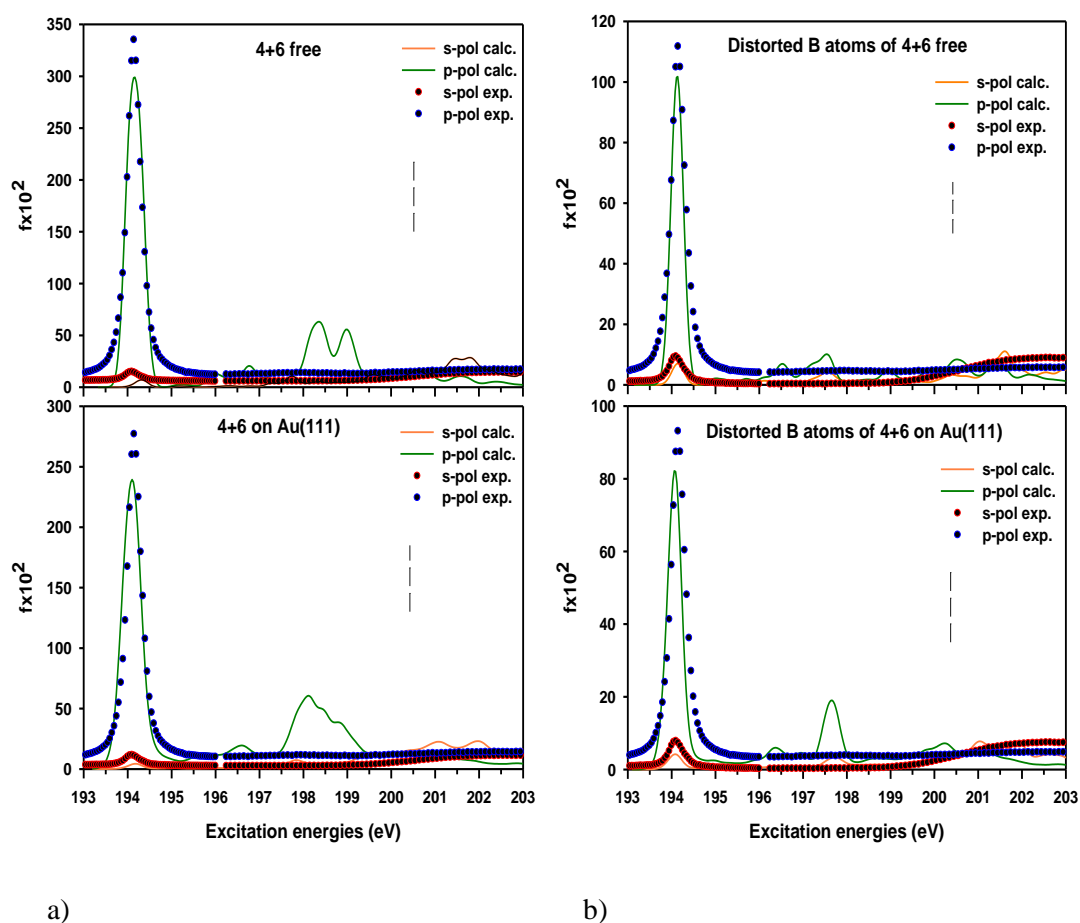
**Figure 23** – Left side: B1s NEXAFS spectra of THDB (upper panel), square (second panel), star (third panel), “4+6” free (fourth panel), “4+6” on Au(111) (lower panel): calculated line-shape (black solid line) with site-resolved B<sub>i</sub> contributions (thick colored vertical bars). Vertical colored dashed bars: ΔKS IPs. Right side: chemical structure of THDB, square, star, “4+6” free and adsorbed on Au(111) with the labeling of the nonequivalent B sites.

From a general overview of Figure 23, one can observe that the progressive presence of a higher number of non-equivalent sites along the series generates the splitting of single non-degenerate transitions occurring in the THDB spectrum into several closely spaced transitions; the relative space of the latter reflects the different chemical environment among the nonequivalent B sites.

Moreover, all NEXAFS spectra are dominated by an intense lower-energy peak (around 191 eV), whose position appears slightly shifted to higher excitation energies (around 191.5 eV) in the THDB spectrum. This structure arises from the B1s transitions to the LUMO, which displays  $\pi$  character localized on the B<sub>i</sub>-B bond. It is interesting to notice that the band extending in the energy range between 194 eV and 196 eV becomes broader along the series of molecules as the number of non-equivalent centers increases; this is also due to the higher molecular strain of the “4+6” system with respect to the star, which derives from a perfectly periodic 2D boroxine network. The molecular strain in the “4+6” system instead arises from the junction of two molecular fragments belonging to different point group symmetry. Also the square system is characterized by a certain strain and distortion from planarity; however, its spectrum does not display significant variations as concerns the intensity and shape of the band located between 194 and 196 eV. Another important consideration regards the fact that, in the “4+6” system, the *distorted* B atoms (*i.e.*, in close proximity of the junction) give a small contribution to the total NEXAFS spectrum in this energy range, pointing toward the presence of defects in the synthesized 2D boroxine-containing COF. The molecular strain in proximity of these atoms induces a rehybridization of the involved MOs, which all exhibit mixed valence-Rydberg character; this justifies the weaker intensity of these transitions with respect to that observed in the NEXAFS spectra of the other investigated systems. Moreover, a smoothening of the intensity profiles is visible in passing from the “4+6” gas-phase system to that one adsorbed on the Au(111) surface. The Au-boroxine interaction, although weak, is also responsible of a redistribution of the NEXAFS spectral intensity of the “4+6” gas-phase system in a large number of transitions; this can be ascribed to a certain degree of orbital hybridization with the substrate’s *sp* states.



A comparison between the experimental 2D COF and theoretical B1s polarized NEXAFS spectra of the “4+6” gas-phase system and adsorbed on the Au(111) surface is displayed in Figure 24a. To support the hypothesis on the presence of defects in the synthesized 2D COF, a further comparison between the experimental 2D COF and computed B1s polarized NEXAFS spectra of the “4+6” gas-phase system and adsorbed on Au(111) has been considered by including only the partial contributions relative to the *distorted* B atoms (see Figure 24b), since these sites feel the presence of the defects to a larger extent. This could approximate a theoretical NEXAFS spectral profile that could simulate the one obtained for a 2D COF overlayer presenting a significant number of defects, and negligible, or at least much less important, contributions from the other B sites.



**Figure 24** – Comparison between experimental (filled circles) and calculated (colored solid lines) B1s (*s*- and *p*-pol) polarized NEXAFS spectra of “4+6” free (upper panel) and adsorbed on Au(111) (lower panel): a) all B atoms, b) only distorted B atoms. Black vertical dashed bars: mean  $\Delta$ KS IPs. The calculated spectra have been shifted by +3.33 eV (a) and +3.20 eV (b) in order to match the first peak of the two experimental profiles.

A general inspection of Figure 24a reveals that the polarized experimental B1s NEXAFS spectra show a pronounced dichroism, ascribable to the flat orientation of the 2D network on the metal surface. It is possible to observe that the theoretical profiles are able to reproduce the large dichroism displayed by intense lower-energy feature. On the other hand, at higher excitation

energies, the experimental spectra display a flat profile, in contrast to the detection of relatively intense structures in the theoretical profiles, around 198-199 eV; this could be associated to the presence of a large number of defects in the synthesized 2D boroxine-containing COF. From an inspection of Figure 24b, a better matching between theory and experiment is found when considering only the partial contributions of the *distorted* B atoms; this is especially evident in the case of the gas phase system, with a considerable decreasing of the intensity of spectral features in the higher-energy region of the calculated spectra. The flat profile observed in the experimental spectra could be also correlated to a slight flattening of the boroxine unit upon the interaction with the Au surface, with an increase of its excess strain energy. It is therefore evident that larger models (*e.g.*, containing 7-fold or 8-fold motifs) could be exploited to provide a more reliable representation of the 2D COF overlayer as well as to obtain a better agreement with the experimental measurements. However, such calculations would be particularly unfeasible due to the need to consider the Au-boroxine interaction.

## Conclusions

In this study, DFT-TP calculations have been performed to simulate the B1s NEXAFS spectra of three possible building block molecules forming a phenyl-free 2D boronic network when THDB is deposited on the Au(111) surface, with reference to a recent work by a group of experimental researchers afferent to the ALOISA beamline of the Elettra Synchrotron in Trieste [14]. The investigated molecules are displayed in Figure 6, and has been denoted as “square”, “star”, and “4+6” systems, in analogy with their structural shape. The aim of this study has been the investigation of the electronic study of the aforementioned building block molecules as well as the choice of the best model to simulate the experimental B1s (*s*- and *p*-pol) polarized NEXAFS spectrum of THDB on the Au(111) surface<sup>14</sup> in order to provide a reliable interpretation of its spectral features. Another objective of the present study has been the determination of the nature of the Au-boroxine interaction; this has been achieved by comparing the B1s NEXAFS spectrum of gas-phase “4+6” system with that relative to the same system adsorbed on the Au(111) surface.

In particular, the analysis of the B1s NEXAFS spectra of the three gas-phase molecules has shown that the “4+6” system represents the best model to reproduce the experimental measurements. However, it has been found out that this model could be further improved by including 7-fold or 8-fold motifs, despite the higher computational effort that would be required for these type of systems. Furthermore, an evolution of the NEXAFS intensity distribution is observed in going from the THDB precursor to the building block molecules due to the presence of a higher

number of non-equivalent B sites as well as to the increase of the molecular strain. Another important finding concerns the small contribution of the *distorted* B atoms of the “4+6” system to the corresponding total B1s NEXAFS spectrum, thus confirming the predominant role of the defects in the synthesized 2D boroxine COF. Finally, it has been possible to conclude that the Au-boroxine interaction is weak (*i.e.*, physisorption), since the spectral profile of the gas-phase “4+6” system substantially resembles that relative to the adsorbate.

## Appendix

**Table 1** – Peak assignments of the total B1s NEXAFS spectrum of the “square” system. Transition energies (eV), oscillator strengths ( $\text{fx}10^2$ ), and the composition of the involved MOs in terms of Mulliken population of AOs centered on each specific atom are reported. Only transitions with  $\text{fx}10^2 \geq 1.00$  are included.

Peaks	Core-hole site	E(eV)	fx10 <sup>2</sup>	Main atomic contribution to the final MO			Assignment
				Bi	B	O	
<b>A</b>	B1	190.91	19.8	47% 2p <sub>z</sub> 2% nd	23% 2p <sub>z</sub>	22% 2p <sub>z</sub>	LUMO/ $\pi$
	B4	190.91	19.6	46% 2p <sub>z</sub> 2% nd	23% 2p <sub>z</sub>	22% 2p <sub>z</sub>	
	B2	191.07	23.7	56% 2p <sub>z</sub>	14% 2p <sub>z</sub>	24% 2p <sub>z</sub> 2% 3d	
	B3	191.07	16.0	39% 2p <sub>z</sub> 3% nd	3% 2p <sub>z</sub>	26% 2p <sub>z</sub>	
<b>B</b>	B3	192.92	2.11	5% 2p <sub>z</sub>	76% 2p <sub>z</sub>	16% 2p <sub>z</sub>	LUMO+3/ $\pi^*$
<b>C</b>	B4	193.58	2.28	6% 2p <sub>z</sub>	75% 2p <sub>z</sub>	15% 2p <sub>z</sub>	LUMO+5/ $\pi^*$
	B1	193.58	2.31	6% 2p <sub>z</sub>	76% 2p <sub>z</sub>	15% 2p <sub>z</sub>	
<b>D</b>	B3	194.03	1.20	54% ns 8% 2s 1% 2p <sub>x</sub>	3% 2s 2% ns	2% 2p <sub>y</sub> 1% 2s	LUMO+7/mixed valence-Rydberg
	B3	194.42	1.89	4% 2p <sub>z</sub>	76% 2p <sub>z</sub>	18% 2p <sub>z</sub>	LUMO+8/ $\pi^*$
	B3	194.58	1.17	47% np 7% 2p <sub>x</sub> 4% ns 1% 4f	1% np	2% 2p <sub>y</sub> 1% 2s	LUMO+9/ mixed valence-Rydberg
<b>E</b>	B3	194.96	14.0	14% 2p <sub>z</sub> 10% np 2% 3d	55% 2p <sub>z</sub> 2% np	13% 2p <sub>z</sub>	LUMO+11/mixed valence-Rydberg
<b>F</b>	B4	195.31	8.58	14% np 10% 2p <sub>z</sub> 4% 3d 2% nd	49% 2p <sub>z</sub> 4% np	11% 2p <sub>z</sub>	LUMO+10/mixed valence-Rydberg
	B1	195.34	10.7	11% 2p <sub>z</sub> 4% 3d 3% np 2% nd	62% 2p <sub>z</sub> 2% np	13% 2p <sub>z</sub>	LUMO+11/mixed valence-Rydberg
	B4	195.36	2.40	46% np 2% 2p <sub>z</sub> 2% 2p <sub>y</sub> 2% ns	19% np 11% 2p <sub>z</sub> 6% 2p <sub>y</sub>	1% 2p <sub>z</sub>	
<b>G</b>	B2	195.71	1.03	22% 3d 5% np	23% np 13% 2p <sub>y</sub>	1% 2p <sub>x</sub>	LUMO+12/ mixed

				4% nd 2% 2p <sub>x</sub> 1% ns			valence-Rydberg
	B1	195.85	1.31	29% 3d 7% np 2% nd 2% 4f	11% np 6% 2p <sub>x</sub> 2% 2p <sub>y</sub> 2% 4f	/	LUMO+14/mixed valence-Rydberg
	B2	195.87	10.5	8% 2p <sub>z</sub> 3% 3d	72% 2p <sub>z</sub>	13% 2p <sub>z</sub>	LUMO+13/mixed valence-Rydberg
<b>H</b>	B1	196.42	1.23	12% nd 3% np 1% 3d 1% ns 1% 2p <sub>y</sub>	21% np 10% 2p <sub>x</sub> 10% 2p <sub>y</sub> 2% 2s	1% 2p <sub>x</sub> 1% 2s	LUMO+21/mixed valence-Rydberg
	B4	196.51	1.43	9% 4f 6% 3d 5% nd 4% np 2% ns 1% 2p <sub>y</sub>	21% np 11% 2p <sub>x</sub> 7% 2p <sub>y</sub>	/	
	B2	196.83	1.33	24% 3d 23% np 3% 2p <sub>y</sub> 2% nd	13% np 6% 2p <sub>x</sub> 4% 2p <sub>y</sub>	/	LUMO+26/mixed valence-Rydberg
	B1	196.85	1.01	24% np 5% nd 5% 2p <sub>y</sub> 3% 3d 2% 4f	22% 2p <sub>z</sub> 2% 2p <sub>y</sub> 1% np	/	LUMO+27/mixed valence-Rydberg

**Table 2** – Peak assignments of the total B1s NEXAFS spectrum of the “star” system. Transition energies (eV), oscillator strengths (fx10<sup>2</sup>), and the composition of the involved MOs in terms of Mulliken population of AOs centered on each specific atom are reported. Only transitions with fx10<sup>2</sup> ≥ 1.00 are included.

Peaks	Core-hole site	E(eV)	fx10 <sup>2</sup>	Main atomic contribution to the final MO			Assignment
				Bi	B	O	
<b>A</b>	B2	190.80	15.2	48% 2p <sub>z</sub> 2% nd	20% 2p <sub>z</sub>	23% 2p <sub>z</sub>	LUMO/π
	B1	190.86	15.3	48% 2p <sub>z</sub> 1% nd	20% 2p <sub>z</sub>	23% 2p <sub>z</sub>	
	B3	190.94	34.6	54% 2p <sub>z</sub> 2% nd	16% 2p <sub>z</sub>	23% 2p <sub>z</sub>	
	B4	191.02	24.1	39% 2p <sub>z</sub> 3% nd	26% 2p <sub>z</sub>	26% 2p <sub>z</sub>	
<b>B</b>	B2	193.10	1.18	2% 2p <sub>z</sub>	80% 2p <sub>z</sub>	15% 2p <sub>z</sub>	LUMO+3/π*
	B4	193.11	2.68	4% 2p <sub>z</sub>	76% 2p <sub>z</sub>	12% 2p <sub>z</sub>	
<b>C</b>	B4	193.50	3.01	4% 2p <sub>z</sub>	76% 2p <sub>z</sub>	16% 2p <sub>z</sub>	LUMO+5/π*
<b>D</b>	B1	193.85	3.01	9% 2p <sub>z</sub>	64% 2p <sub>z</sub>	18% 2p <sub>z</sub>	LUMO+6/π*
<b>E</b>	B3	194.39	1.11	87% ns 1% np <sub>y</sub> 1% nd	4% 2p <sub>y</sub> 3% ns 2% np <sub>y</sub>	/	LUMO+7/mixed valence-Rydberg
	B4	194.52	1.92	5% 2p <sub>z</sub>	73% 2p <sub>z</sub>	17% 2p <sub>z</sub>	LUMO+9/π*

	B2	194.54	1.18	4% 2p <sub>z</sub>	75% 2p <sub>z</sub>	18% 2p <sub>z</sub>	
<b>F</b>	B4	194.99	20.7	12% 2p <sub>z</sub> 15% np <sub>z</sub> 3% 3d	52% 2p <sub>z</sub> 3% np <sub>z</sub>	13% 2p <sub>z</sub>	LUMO+12/ $\pi^*$
	B2	195.28	8.09	14% 2p <sub>z</sub> 5% 3d 2% np <sub>z</sub> 1% nd	61% 2p <sub>z</sub> 3% np <sub>z</sub>	13% 2p <sub>z</sub>	
	B1	195.43	7.53	11% 2p <sub>z</sub> 7% nd	62% 2p <sub>z</sub> 4% np <sub>z</sub>	13% 2p <sub>z</sub>	
<b>G</b>	B3	195.61	1.36	25% 3d 5% nd 2% ns 2% 2p <sub>x</sub> 1% np <sub>y</sub>	3% 2p <sub>x</sub> 3% np <sub>y</sub> 3% np <sub>x</sub>	1% np <sub>y</sub>	LUMO+13/mixed valence-Rydberg
	B3	195.68	16.7	7% 2p <sub>z</sub> 6% np <sub>z</sub> 1% 3d 1% nd	67% 2p <sub>z</sub>	13% 2p <sub>z</sub>	LUMO+15/ $\pi^*$
<b>H</b>	B2	195.96	1.20	10% np <sub>y</sub> 6% 2p <sub>y</sub> 3% 3d 2% np <sub>x</sub> 1% 2p <sub>x</sub>	14% np 11% 2p <sub>x</sub> 3% 2s 3% nd 3% nf 1% 2p <sub>y</sub>	3% 2p <sub>y</sub>	LUMO+18/mixed valence-Rydberg
	B1	196.97	1.13	16% np 9% nd 4% nf 2% 2p <sub>x</sub>	4% 2p <sub>x</sub> 2% np <sub>y</sub> 1% 2s	2% 2p <sub>y</sub>	LUMO+31/mixed valence-Rydberg

**Table 3** – Peak assignments of the total B1s NEXAFS spectrum of the “4+6” gas-phase system. Transition energies (eV), oscillator strengths ( $\text{fx}10^2$ ), and the composition of the involved MOs in terms of Mulliken population of AOs centered on each specific atom are reported. Only transitions with  $\text{fx}10^2 \geq 0.40$  are included.

Peaks	Core-hole site	E(eV)	$\text{fx}10^2$	Main atomic contribution to the final MO			Assignment
				Bi	B	O	
<b>A</b>	B52	190.67	10.2	48% 2p <sub>z</sub> 2% nd	19% 2p <sub>z</sub>	23% 2p <sub>z</sub>	
	B23	190.68	10.5	49% 2p <sub>z</sub>	18% 2p <sub>z</sub>	20% 2p <sub>z</sub>	
	B56	190.69	10.1	47% 2p <sub>z</sub> 1% nd	20% 2p <sub>z</sub>	22% 2p <sub>z</sub>	
	B50	190.71	10.2	48% 2p <sub>z</sub> 1% nd	19% 2p <sub>z</sub>	22% 2p <sub>z</sub>	
	B57	190.71	10.0	47% 2p <sub>z</sub> 1% nd	21% 2p <sub>z</sub>	21% 2p <sub>z</sub>	
	B51	190.79	11.4	53% 2p <sub>z</sub> 1% nd	17% 2p <sub>z</sub>	23% 2p <sub>z</sub> 1% 3d	

	B17	190.80	11.8	55% 2p <sub>z</sub>	14% 2p <sub>z</sub>	22% 2p <sub>z</sub> 2% 3d	LUMO/ $\pi$
	B58	190.85	11.5	53% 2p <sub>z</sub> 2% nd	17% 2p <sub>z</sub>	22% 2p <sub>z</sub>	
	B18	190.89	10.9	47% 2p <sub>z</sub> 2% nd	18% 2p <sub>z</sub>	17% 2p <sub>z</sub> 1% 2p <sub>x</sub>	
	B25	190.92	7.91	38% 2p <sub>z</sub> 3% 3d	27% 2p <sub>z</sub>	25% 2p <sub>z</sub>	
	B64	190.93	7.93	38% 2p <sub>z</sub> 3% nd	26% 2p <sub>z</sub>	26% 2p <sub>z</sub>	
	B59	190.96	7.90	38% 2p <sub>z</sub> 3% nd	26% 2p <sub>z</sub>	26% 2p <sub>z</sub>	
	B24	190.98	11.0	44% 2p <sub>z</sub> 5% 2p <sub>y</sub> 1% nd	19% 2p <sub>z</sub>	15% 2p <sub>z</sub>	
	B16	190.99	11.2	47% 2p <sub>z</sub> 4% 2p <sub>y</sub> 2% nd	18% 2p <sub>z</sub>	16% 2p <sub>z</sub>	
	B22	191.06	11.7	40% 2p <sub>z</sub> 8% 2p <sub>x</sub>	23% 2p <sub>z</sub>	11% 2p <sub>z</sub> 3% 2p <sub>x</sub>	
<b>B</b>	B59	192.58	0.67	3% 2p <sub>z</sub>	75% 2p <sub>z</sub>	14% 2p <sub>z</sub>	LUMO+2/ $\pi^*$
	B25	192.59	0.41	2% 2p <sub>z</sub>	76% 2p <sub>z</sub>	8% 2p <sub>z</sub>	
	B64	192.63	0.68	3% 2p <sub>z</sub>	75% 2p <sub>z</sub>	12% 2p <sub>z</sub>	
<b>C</b>	B24	193.28	0.95	2% 2p <sub>z</sub>	53% 2p <sub>z</sub>	11% 2p <sub>z</sub>	LUMO+7/ $\pi^*$
	B18	193.39	0.96	3% 2p <sub>z</sub>	69% 2p <sub>z</sub>	12% 2p <sub>z</sub>	LUMO+8/ $\pi^*$
	B23	193.52	0.98	5% 2p <sub>z</sub>	71% 2p <sub>z</sub>	13% 2p <sub>z</sub>	LUMO+9/ $\pi^*$
<b>D</b>	B22	193.95	1.93	2% 2p <sub>z</sub> 4% 2p <sub>x</sub>	64% 2p <sub>z</sub> 4% 2p <sub>x</sub>	7% 2p <sub>z</sub>	LUMO+12/valence
	B16	194.10	1.37	3% 2p <sub>z</sub> 3% 2p <sub>y</sub>	62% 2p <sub>z</sub> 4% 2p <sub>y</sub>	7% 2p <sub>z</sub>	LUMO+11/valence
	B16	194.23	0.79	2% 2p <sub>z</sub> 2% 2p <sub>y</sub> 1% ns	64% 2p <sub>z</sub> 2% 2p <sub>y</sub>	9% 2p <sub>z</sub>	LUMO+12/valence
	B24	194.40	1.51	4% 2p <sub>y</sub> 2% ns	51% 2p <sub>z</sub> 6% 2p <sub>x</sub> 4% 2p <sub>y</sub>	9% 2p <sub>z</sub>	LUMO+14/mixed valence-Rydberg
	B18	194.43	2.18	5% 2p <sub>x</sub> 5% ns 2% 2p <sub>z</sub>	47% 2p <sub>z</sub> 4% 2p <sub>x</sub>	6% 2p <sub>z</sub>	LUMO+13/mixed valence-Rydberg
<b>E</b>	B25	194.75	3.31	8% 2p <sub>z</sub> 1% np	65% 2p <sub>z</sub> 1% 2p <sub>y</sub>	10% 2p <sub>z</sub>	LUMO+16/valence
	B25	194.82	2.98	6% 2p <sub>z</sub> 5% np	48% 2p <sub>z</sub>	6% 2p <sub>z</sub>	LUMO+17/mixed valence-Rydberg
	B59	194.85	1.24	40% np 12% ns 8% 3d 6% nd 5% 2p <sub>x</sub> 3% 2p <sub>z</sub>	5% 2p <sub>z</sub> 4% 2p <sub>x</sub> 1% ns	2% 2p <sub>y</sub>	
	B64	194.87	6.26	15% np 11% 2p <sub>z</sub> 2% 3d	48% 2p <sub>z</sub> 3% np	11% 2p <sub>z</sub>	LUMO+18/mixed valence-Rydberg
	B59	194.89	6.25	13% np 11% 2p <sub>z</sub> 2% 3d 2% 2p <sub>x</sub> 2% ns	49% 2p <sub>z</sub> 2% np	11% 2p <sub>z</sub>	

	B57	194.99	3.78	11% 2p <sub>z</sub> 6% np 2% ns 1% 3d	53% 2p <sub>z</sub> 3% 2p <sub>x</sub> 1% 2s	8% 2p <sub>z</sub>	LUMO+16/mixed valence-Rydberg
	B23	195.00	2.24	7% ns 5% 3d 3% 2p <sub>z</sub> 2% np	44% 2p <sub>z</sub> 7% 2p <sub>x</sub> 2% 2s	6% 2p <sub>z</sub>	LUMO+17/mixed valence-Rydberg
	B52	195.06	2.87	22% np 6% 2p <sub>z</sub> 4% 2p <sub>x</sub> 2% 3d 2% ns	32% 2p <sub>z</sub> 4% 2p <sub>x</sub> 2% 2p <sub>y</sub> 2% np	6% 2p <sub>z</sub>	
	B56	195.09	3.12	38% np 4% 2p <sub>z</sub> 3% 3d 2% ns 1% nd	38% 2p <sub>z</sub> 3% np 1% 2p <sub>x</sub> 1% 2p <sub>y</sub> 1% 2s	7% 2p <sub>z</sub>	
	B50	195.13	4.13	11% np 10% 2p <sub>z</sub> 3% 2p <sub>y</sub> 2% 3d	47% 2p <sub>z</sub> 1% np	10% 2p <sub>z</sub>	
	B52	195.17	2.19	39% np 4% 2p <sub>z</sub> 4% 2p <sub>x</sub> 2% ns 1% nd	26% 2p <sub>z</sub> 2% np	4% 2p <sub>z</sub>	
	B56	195.20	1.85	39% np 4% 2p <sub>z</sub> 3% 3d 2% ns 1% nd	21% 2p <sub>z</sub> 5% np 1% 2p <sub>x</sub>	3% 2p <sub>z</sub>	
<b>F</b>	B23	195.45	1.81	22% 3d 7% ns 6% 2p <sub>z</sub> 4% np	21% 2p <sub>z</sub> 10% np 6% 2p <sub>y</sub>	6% 2p <sub>z</sub>	LUMO+21/mixed valence-Rydberg
	B58	195.60	1.47	19% np 7% 3d 1% ns	35% 2p <sub>z</sub> 10% np 8% 2p <sub>y</sub> 3% 2p <sub>x</sub>	2% 2p <sub>z</sub>	
	B51	195.62	4.94	16% np 4% 2p <sub>z</sub>	60% 2p <sub>z</sub> 2% np	10% 2p <sub>z</sub>	
	B17	195.69	1.22	7% 3d 5% 2p <sub>z</sub> 3% np	40% np 18% 2p <sub>z</sub> 4% 2p <sub>x</sub> 4% 2p <sub>y</sub>	4% 2p <sub>z</sub>	LUMO+22/mixed valence-Rydberg
	B58	195.73	2.64	15% 2p <sub>z</sub> 14% np 7% 3d 7% nd	47% 2p <sub>z</sub>	11% 2p <sub>z</sub>	LUMO+23/mixed valence-Rydberg
	B17	195.80	1.51	37% np 9% ns 4% 3d	14% 2p <sub>z</sub> 8% np 5% 2p <sub>x</sub> 3% 2p <sub>y</sub>	2% 2p <sub>z</sub>	LUMO+24/mixed valence-Rydberg

## Bibliography

---

- <sup>1</sup> Y. Shen, A.R. Hosseini, M. H. Wong, G.G. Malliaras. *Chem. Phys. Chem.* **2004**, *5*, 16-25
- <sup>2</sup> A. Kahn, N. Koch, W. Gao., *Pol. Phys.* **2003**, *41*, 2529-2548
- <sup>3</sup> W. Auwärter, A. Schriffin, A. Weber-Bargioni, Y. Pennec, A. Riemann, *Intern. J. of Nanotech.* **2008**, *5*, 1171-1193
- <sup>4</sup> J.V. Barth, G. Costantini, K. Kern., *Nature* **2005**, *437*, 671-679
- <sup>5</sup> M. Treier, N.V. Richardson, R. Fasel, *J. Am. Chem. Soc.* **2008**, *130*, 14054–14055
- <sup>6</sup> S.W. Hla, L. Bartels, G. Meyer, K.H. Rieder, *Phys. Rev. Lett.* **2000**, *85*, 2777–2780
- <sup>7</sup> M. Di Giovannantonio, M. El Garah, J. Lipton-Duffin, V. Meunier, L. Cardenas, Y. Fagot Revurat, A. Cossaro, A. Verdini, D.F. Perepichka, F. Rosei, G. Contini, *ACS Nano* **2013**, *7*, 8190–8198
- <sup>8</sup> A.C. Marele, R. Mas-Balleste, L. Terracciano, J. Rodriuez-Fernadez, I. Berlanga, S. S. Alexandre, R. Otero, J. Gallego, F. Zamora, J. Goetz-Rodriuez, *Chem. Commun.* **2012**, *48*, 6749–6864
- <sup>9</sup> R. Nishiyabu, Y. Kubo, T.D. James, J.S. Fossey, *Chem. Commun.* **2011**, *47*, 1124–1150
- <sup>10</sup> R. Wang, X.-R. Zhang, S. Wang, G. Fu, J. Wang, *Phys. Chem. Chem. Phys.* **2015**, *18*, 1258–1264
- <sup>11</sup> J. Teyssandier, S. De Feyter, *Chem. Commun.* **2016**, *52(77)*, 11465-11487
- <sup>12</sup> D. Toffoli, M. Stredansky, Z. Feng, G. Balducci, S. Furlan, M. Stener, H. Ustunel, D. Cvetko, G. Kladnik, A. Morgante, A. Verdini, C. Dri, G. Comelli, G. Fronzoni, A. Cossaro, *Chem. Sci.* **2017**, *8*, 3789-3798
- <sup>13</sup> R.-N. Wang, X. Zhang, S.-F. Wang, G. Fu, J. Wang, *Phys. Chem. Chem. Phys.* **2016**, *18*, 1258–1264
- <sup>14</sup> M. Stredansky, A. Sala, T. Fontanot, R. Costantini, C. Africh, G. Comelli, L. Floreano, A. Morgante, A. Cossaro, *Chem. Commun.* **2018**, *54*, 3971-3973
- <sup>15</sup> P. Giannozzi *et al.*, *J. Phys.: Condens. Matter*, **2009**, *21*, 395502–395520
- <sup>16</sup> R.G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, **1989**
- <sup>17</sup> D. Vanderbilt, *Phys. Rev. B: Condens. Matter Mater. Phys.*, **1990**, *41*, 7892–7895
- <sup>18</sup> J.P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.*, **1996**, *77*, 3865–3868
- <sup>19</sup> H. Rydberg, N. Jacobson, P. Hyldgaard, S.I. Simak, B.I. Lundqvist, D.C. Langreth, in *Surface Science*, **2003**, vol. 532–535, 606–610
- <sup>20</sup> C.G.J. Broyden, *Inst. Math. Its Appl.*, **1970**, *6*, 76–231; R. Fletcher, *Comput. J.*, **1970**, *13*, 317–322; D. Goldfarb, *Math. Comput.*, **1970**, *24*, 23-26; D.F. Shanno, *Math. Comput.*, **1970**, *24*, 647-656



- 
- <sup>21</sup> C. Fonseca Guerra, J.G. Snijders, G.te Velde, E.J. Baerends, *Theor. Chem. Acc.* **1998**, *99*, 391-403; E.J. Baerends,; D.E. Ellis, P. Roos, *Chem. Phys.* **1973**, *2*, 41-51
- <sup>22</sup> J.P. Perdew, *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *33*, 8822–8824
- <sup>23</sup> J.C. Slater, *Adv. Quantum Chem.* **1972**, *6*, 1–92; L. Triguero, L.G.M. Pettersson, H. Ågren, *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *58*, 8097–8110
- <sup>24</sup> M. Romeo, G. Balducci, M. Stener, G. Fronzoni, *J. Phys. Chem. C*, **2014**, *118*, 1049–1061
- <sup>25</sup> G. Fronzoni, G. Balducci, R. De Francesco, M. Romeo, M. Stener, *J. Phys. Chem. C*, **2012**, *116*, 18910–18919
- <sup>26</sup> G. Balducci, M. Romeo, M. Stener, G. Fronzoni, D. Cvetko, A. Cossaro, M. Dell’Angela, G. Kladnik, L. Venkataraman, A. Morgante, *J. Phys. Chem. C*, **2015**, *119*, 1988–1995
- <sup>27</sup> C. Dri, G. Fronzoni, G. Balducci, S. Furlan, M. Stener, Z. Feng, G. Comelli, C. Castellarin-Cudia, D. Cvetko, G. Kladnik, A. Verdini, L. Floreano, A. Cossaro, *J. Phys. Chem. C*, **2016**, *120*, 6104–6115
- <sup>28</sup> N. Danielis, *Spettroscopie di core per l’analisi della struttura elettronica di network boronici su superfici metalliche*, Bachelor’s Thesis, University of Trieste, **2017**

## Conclusions and perspectives

The present PhD thesis deals with the simulation and interpretation of data obtained from *X-Ray Photoelectron Spectroscopy* (XPS) and *Near Edge X-ray Absorption Fine Structure Spectroscopy* (NEXAFS) of molecules in gas phase and adsorbed on metal surfaces. In particular, studies in the gas phase aimed at achieving a detailed knowledge of the electronic structure of a series of molecular building blocks with promising applications in the design of novel high-performance electronic and optoelectronic devices, whereas investigations on surface chemistry have focused on the elucidation of the electronic structure of boron-containing adsorbates suitable for surface functionalization. Theoretical calculations by means of Density Functional Theory (DFT) and its linear-response time-dependent formulation (TDDFT) have been performed to provide a solid support to XPS and NEXAFS experimental measurements.

The research activity has been firstly devoted to three heterocycles playing fundamental roles in chemistry and biology, namely, the indole molecule and its 2,3-dihydro-7-azaindole (7-AI) and 3-formylindole (3-FI) derivatives. Further investigations have been focused on a series of gas phase  $\pi$ -conjugated molecular complexes with potential applications in the field of organic molecular electronics. The studied material has been the ambipolar 2,8-bis(diphenylphosphoryl)-dibenzo[b,d]thiophene (PPT), whose electronic structure has been explored starting from a detailed investigation on that relative to its building block moieties. A first study has involved the small thiophene precursor and its benzo-annulated derivatives, namely, benzo[b]thiophene (BBT) and DBT, while further investigations have focused on DBT and triphenylphosphine oxide (TPPO). Afterwards, the research activity has dealt with the titanyl phthalocyanine (TiOPc) and its interaction with the Ag(111) surface. A series of important building blocks of Covalent Organic Frameworks (COFs), namely, boric acid [B(OH)<sub>3</sub>], phenyl boronic acid (PBA), and 1,4-phenyl diboronic acid (PDBA), has been the object of the last part of the three year period of PhD, while final studies aimed at determining the nature of the interaction between a series of boronic networks and the Au(111) surface.

The employed computational methods have provided valuable and original outcomes, with new insights on the nature of the virtual states implicated in the X-ray absorption process. Furthermore, in many instances, the agreement between theory and experiment has proven to be remarkable, allowing a reliable assignment of the main spectral features, and their relative association with the different atomic contributions. In particular, the DFT-TP (*Transition Potential*) method in the half core hole approximation (HCH) has turned out to be a powerful approach for the description of the *K-edge* NEXAFS spectra, both in terms of energy separation and intensity distribution of the

observed spectral features. However, in the case of the investigated gas phase boronic systems, we found out that the DFT-TP approach correctly reproduces the B *K-edge* NEXAFS spectrum of the simplest B(OH)<sub>3</sub> molecule, but is not adequate for the corresponding ones relative to PBA and PDDBA, independently of the exchange-correlation functional employed. A satisfactory agreement with the experimental measurements is achieved by including static correlation effects through explicitly correlated wave function multiconfigurational self-consistent field (MCSCF) methods. On the other hand, the DFT-TP approach has proven to be very efficient in the calculation of the C *K-edge* NEXAFS spectrum of PBA; evidently, the relaxation of the charge density upon the core-hole formation is less pronounced in the case of the shallower B1s core state, compared to C, N, or O core excited states, and the ability to accurately describe the residual electron correlation effects is crucial to obtain a satisfactory agreement with the experimental data. On the other hand, the relativistic TDDFT method is able to accurately reproduce the experimental S,P *L<sub>II,III</sub>-edge* NEXAFS spectra of the investigated systems by including the configuration mixing among different excitation channels from degenerate core-holes. Finally, it is worth noting that, from the computational point of view, DFT-TP and TDDFT are the only applicable methods for the simulation of core-electron excitation spectra of molecules adsorbed on metal surfaces. In particular, when applied in conjunction with core absorption spectroscopy, these methods are able to provide interesting information on the adsorption geometry and bonding of molecules on metal surfaces as well as on the adsorbate-substrate interaction<sup>1,2</sup>.

Future investigations will be focused on the calculation of vibrationally resolved NEXAFS spectra of gas phase systems, with the aim to evaluate the influence of vibronic effects, which are known to be able to modulate the intensity pattern and band-shape<sup>3-5</sup>. The research activity could be also further developed to allow for the simulation of pump-probe NEXAFS experiments of both gas phase molecules and adsorbates, which would provide useful insights into valence-excited states, and further stimulate experimental work on time-resolved core-electron spectroscopies<sup>6,7</sup>.

## References

---

- <sup>1</sup> M. Romeo, G. Balducci, M. Stener, G. Fronzoni, *J. Phys. Chem. C*, **2014**, *118*, 1049–1061
- <sup>2</sup> R. De Francesco, M. Stener, G. Fronzoni, *J. Phys. Chem.* **2009**, *11*, 1146–1151
- <sup>3</sup> G. Fronzoni, O. Baseggio, M. Stener, W. Hua, G. Tian, Y. Luo, B. Apicella, M. Alfé, M. de Simone, A. Kivimäki, M. Coreno, *J. Chem. Phys.* **2014**, *141*, 044313-1; 044313-9
- <sup>4</sup> A. Baiardi, M. Mendolicchio, V. Barone, G. Fronzoni, G.A. Cardenas Jimenez, M. Stener, C. Grazioli, M. de Simone, M. Coreno, *J. Chem. Phys.* **2015**, *143*, 204102-1; 204102-12
- <sup>5</sup> M. Mendolicchio, A. Baiardi, G. Fronzoni, M. Stener, C. Grazioli, M. de Simone, V. Barone, *J. Chem. Phys.* **2019**, *151*, 124105-1; 124105-8
- <sup>6</sup> C. Ehlert, M. Gühr, P. Saalfrank, *J. Chem. Phys.* **2018**, *149*, 144112-1; 144112-13
- <sup>7</sup> T. J.A. Wolf et al., *Nat. Commun.* **2016**, *8*:29, 1-7

## **Acknowledgements**

First and foremost, I would like to express my sincere gratitude to my supervisor, prof. Daniele Toffoli, whose great expertise and precious support, both emotional and technical, have strongly motivated me, not only during this three-year period of PhD, but also during the whole Academic career. Furthermore, I would like to thank him for his constant understanding and patience, but mostly for having encouraged me to believe in myself.

My sincere thanks also go to prof. Giovanna Fronzoni, for her knowledge and helpfulness, but in particular for her kindness and encouragement during my Academic career.

I would like to acknowledge Dr. Oscar Baseggio, Dr. Aurora Ponzi, and Dr. Nicola Quadri from the University of Trieste, for hosting me in their office during the first period of my PhD studies as well as for offering me a fundamental help whenever I needed for. Special thanks also go to my PhD colleague, dott. Marco Medves, for his helpfulness and great support.

Last but not the least, I would like to thank my special supporter, Nikola, for sticking with me in the most difficult moments. I would also express my heartfelt thanks to my family, for the precious economical and emotional support they provided me not only during my Academic career, but also during my entire life. In particular, I would like to thank my grand-mother, Zdenka, and my uncle, Davide, who have always encouraged and believed in me, helping me in realizing my dreams and achieving this important goal.