Vibronic fine structure in the nitrogen 1*s* photoelectron spectra from Franck-Condon simulation. II. Indoles

Minrui Wei [1], Lu Zhang,[1] Guangjun Tian,[2] and Weijie Hua [1,*]

[1]*MIIT Key Laboratory of Semiconductor Microstructure and Quantum Sensing, Department of Applied Physics, School of Science, Nanjing University of Science and Technology, 210094 Nanjing, China*

[2]*Key Laboratory for Microstructural Material Physics of Hebei Province, School of Science, Yanshan University, 066004 Qinhuangdao, China*

The vibronic coupling effect in nitrogen 1*s* x-ray photoelectron spectra was systematically studied for a family of 17 bicyclic indole molecules by combining Franck-Condon simulations (including the Duschinsky rotation effect) and density-functional theory. The simulated vibrationally resolved spectra of four molecules agree well with available experiments. Reliable predictions for this family further allows us to summarize rules for spectral evolution in response to three types of common structural changes (side chain substitution, CH ↔ N replacement, and isomerization). Interestingly, vibronic properties of amine and imine nitrogen are clearly separated. They show negative and positive zero-point vibration energies of the core-ionized state with respect to the ground state, respectively, indicating flatter and steeper potential energy surfaces induced by the N 1*s* ionization; amine N atoms show stronger mode mixing effects than imine N atoms; and the 1*s* ionizations on two types of nitrogen led to distinct changes in local bond lengths and angles. The rules are useful for a basic understanding of vibronic coupling in this family, and the precise spectra are useful for future reference and data mining studies.

I. INTRODUCTION

X-ray photoelectron spectroscopy (XPS) is widely used to characterize molecular and material structures nowadays. The core-level binding energy (BE) is both element selective and sensitive to local bonding. For gas molecules, high-resolution vibrationally resolved XPS spectra provide richer information beyond binding energies: the identifying characteristics reflect information about the potential energy surfaces (PESs) of the initial (ground) and final (core-ionized) states, as well as the electronic and nuclear dynamics during the core-ionization process [1–7].

The groundbreaking work of Sieghabn *et al.* [8–10] has significantly improved the precision of XPS (i.e., electronic spectroscopy for chemical analysis, as coined therein). Since then, a vast amount of XPS data have been collected. The technique has enabled the characterization of the structures of many molecules and materials and has helped scientists understand molecular physics, guide the material design, and so on. In such a characterization process, spectral interpretation is an essential final step that translates the detected numerical data into physical and chemical insights. Various experimental databases [11–16] were constructed to assist the interpretation process, with the most famous being the National Institute of Standards and Technology database [11].

However, existing experimental XPS databases have two major drawbacks: variations in BEs among different experiments and the lack of profile data. Different experiments of the same compound have reported discrepant BEs, sometimes over 1 eV [11,17]. The discrepancy mainly comes from the calibration process, which can be highly arbitrary [18]. Meanwhile, most databases collect only BE values, without including the spectral profiles. Efforts have also been made to build online databases [14–16] where raw pictures of the original experimental spectra, with both the BE and profile information, were collected, which facilitated the comparison process for various different compounds. However, the available information is still limited and possible inconsistency in different experiments can impede close analysis.

For these two reasons, constructing a theoretical library for high-resolution XPS is necessary. Data achieved on the same footing provide fair accuracy to analyze the physical rules behind the data, especially among structurally similar systems. Before generating a huge amount of data, at the current stage, it is more important to guarantee reliable data. We wish to validate and optimize the simulation procedure and try to investigate general rules for vibronic coupling based on limited molecules with specific structural similarities. Our initial studies have shown great agreement with the experiments by combing the full core hole (FCH) density-functional theory (DFT) and the Franck-Condon simulations [17,19,20]. This motivates us to expand the use of this protocol to wider systems and gain more comprehensive insights into vibronic coupling within a family. Quantitative assessment of the impact of structural changes on molecular properties is meaningful for further understanding the structure-spectroscopy relation, which would provide insight for future data mining studies.
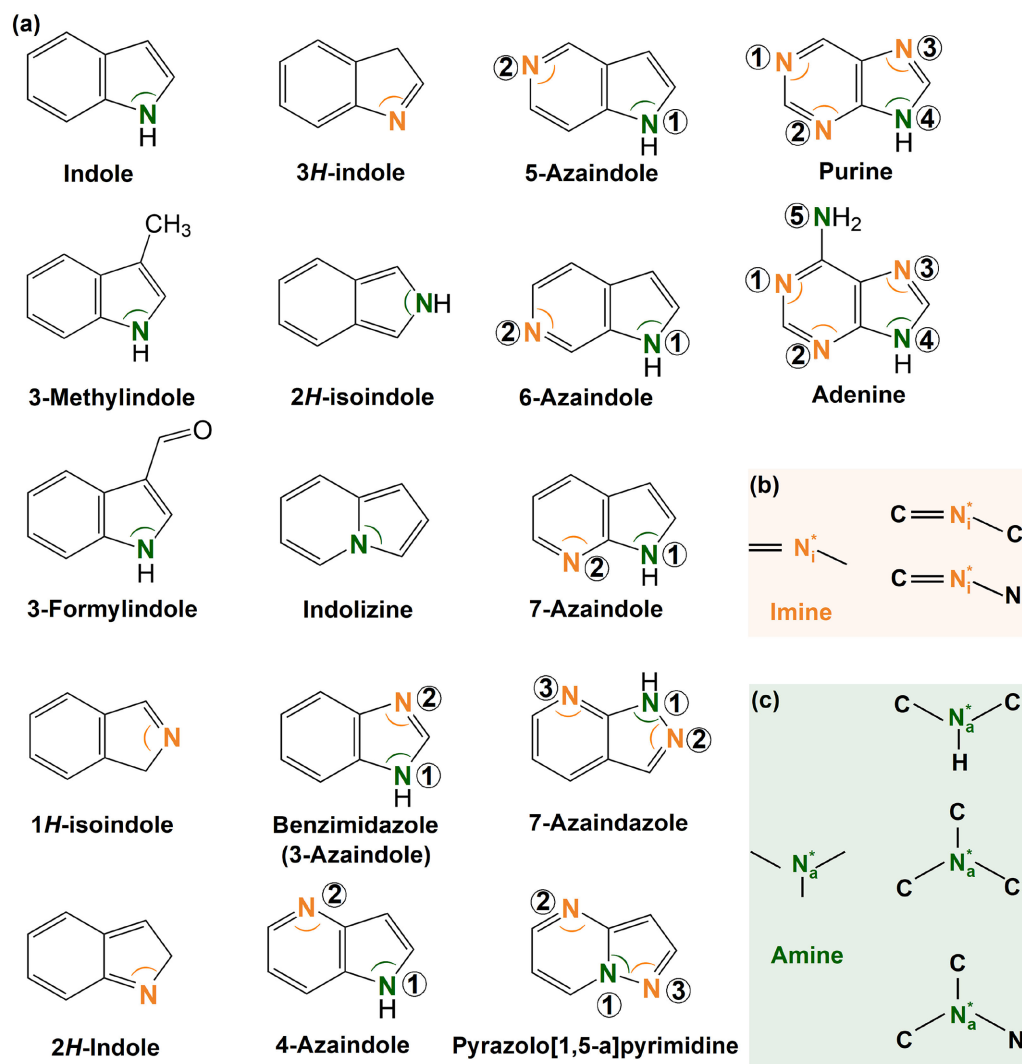
_____

*wjhua@njust.edu.cn

FIG. 1. (a) Structures of all 17 molecules under study, aligned in column-major order with increasing (1, 2, . . . , 5) N atoms. In molecules with more than one N atom, all N atoms are indexed by numbers in circles (not to be confused with the atomic indices as used in nomenclature). Amine (green) and imine (orange) N atoms are distinguished by colors. Selected angles are labeled by arcs. (b) Two imine-type and (c) three amine-type local structures at the ionized N center (denoted by N_i^* and N_a^*, respectively) extracted from the 17 molecules.

Aromatic nitrogen-containing heterocyclic molecules (N-heterocycles) [21] serve as an ideal family for testing new theoretical methods [22], which are also important building blocks for biomolecules [23], high-energy-density compounds [24], nonlinear optical materials [25], and pharmacologically active compounds [26]. In paper I [17] of this series, our calculations on pyrimidine showed good agreement with the high-resolution gas-phase experiment [27]; accurate vibrationally resolved N 1s XPS spectra for a group of azine molecules were predicted and analyzed to understand the rules for vibronic coupling effects as influenced by consecutive replacement of the CH group with an N atom.

As a continuation, in this work (paper II) we further investigate the indole family. Indole has long been recognized as a star molecule in the field of biochemistry, being a structural motif for numerous drugs and natural products [28–32]. Structurally, azines are six-membered rings and indoles are fused by one six- and one five-membered rings. Predictions for five-membered ring compounds (where there are fewer

experimental data) are left for future work. We hope our work can construct a complete picture to understand the N 1s vibronic fine structure for small aromatic N-heterocycles.

Figure 1(a) depicts the 17 selected common indole-derived molecules. Our choice of sample set covers three common types of structural changes: (1) isomerization, (2) side chain substitution, and (3) CH ↔ N replacement. We choose six indole isomers in this study, where three contain imine (=N–) and three contain amine (–N <) nitrogen atoms [Figs. 1(b) and 1(c)]. This sample set is to examine the structure-spectroscopy relation in response to the local bonding change at the ionization center N^* (N_i^* or N_a^* for imine or amine nitrogen, respectively). Significant differences between the two types were found for N 1s binding energies, x-ray absorption spectra (XAS), and resonant inelastic x-ray scattering (RIXS) spectra of small molecules [33,34], large DNA double strands [34], and two-dimensional material g-C_3N_4 [35]. We will investigate the XPS fine structures and vibronic properties as related to the two types.

Side chain substitution can push the electrons to or pull the electrons off the indole ring. Two substituted derivatives with available experimental spectra are selected in our study: one –CH$_3$ (3-methylindole) and one –CHO (3-formylindole) substitute. The –CH$_3$ is a weak electron-donating group (EDG) and –CHO is a moderate electron-withdrawing group. The two molecules are important precursors for the synthesis of some biologically or pharmacologically active compounds [36].

Meanwhile, nine molecules are chosen to demonstrate the different degrees (1, 2, and 3) of CH $\leftrightarrow$ N replacement. Various azaindoles represent an important class of organic molecules with one N substitution on the indole ring, which are fine chemical intermediates commonly used in medicine [37], biological materials [38], and natural products [39]. A total of five azaindoles are selected, among which benzimidazole (also named 3-azaindole) is a special one where two nitrogen atoms are both located in the five-membered ring (cf. the rest, with one nitrogen atom on one ring). In addition, two molecules each for double (7-azaindazole and pyrazolo[1,5-a]pyrimidine) and triple (purine and adenine) CH $\leftrightarrow$ N replacements are selected. Structurally, an adenine is simply an –NH$_2$ substituted purine. Adenine is also one of the building block molecules for DNA, responsible for the radiation loss and photolysis properties.

Experimental N 1s XPS spectra are only available for indole [40], 3-methylindole [41], 3-formylindole [40], and adenine [4] among all systems under study. The spectral resolutions are not always high enough to see clear vibronic structures (especially for indole and 3-formylindole [40]), but the peak asymmetry is obvious for every molecule, indicating significant influences of the vibronic coupling effects. On the other hand, theoretical XPS studies were done for selected systems, but were limited to pure vertical excitations and electronic-only calculations. For example, recently, He *et al.* [42] simulated the N 1s XPS spectra of indole at the second-order Møller-Plesset perturbation theory (MP2) level with a relatively large half-width at half maximum (HWHM) of 0.58 eV and achieved a generally similar profile to the experiment. Plekan *et al.* [4] and Wang *et al.* [43] simulated the XPS spectra of adenine by DFT with relatively large shifts in absolute binding energies (-1.32 [4] and 1.65 eV [43]). The goal of this study is to provide high-precision vibrationally resolved XPS spectra, based on which to carry out systematic analyses (dominant vibronic transitions, active vibrational modes, contributions of 0-n transitions, structural changes induced by core ionization, mode mixing, etc.) and yield general rules on the vibronic coupling.

II. COMPUTATIONAL METHODS

Details were presented in paper I [17]. Briefly, all electronic structure calculations were first performed by using GAMESS-US [44,45], where the DFT method with the Becke three-parameter Lee-Yang-Par (B3LYP) functional [46] was employed. Then Franck-Condon simulations including the Duschinsky rotation (DR) [47] effects were used to generate the vibronic fine structures by using the modified [19] DYNAVIB package [48]. All ground-state (GS) calculations were performed using the cc-PVTZ basis set [49,50]. A

consistent basis set was adopted for the core-ionized state and a double basis set technique [17,19] was used. Vertical and adiabatic ionic potentials (IPs) and the 0-0 vibrational transition energy were computed, respectively, via [17,19]

$$I^{\text{vert}} = E_{\text{FCH}}|_{\text{min GS}} - E_{\text{GS}}|_{\text{min GS}} + \delta_{\text{rel}}, \quad (1)$$

$$I^{\text{ad}} = E_{\text{FCH}}|_{\text{min FCH}} - E_{\text{GS}}|_{\text{min GS}} + \delta_{\text{rel}}, \quad (2)$$

$$E_{00}^{\text{DR}} = I^{\text{ad}} + \Delta\varepsilon_0. \quad (3)$$

Here E_{GS} (E_{FCH}) denotes the total energy of the GS (FCH) state and min GS (min FCH) represents the optimized geometry of the GS (FCH) state. In addition, $\Delta\varepsilon_0$ is the difference of the zero-point vibrational energies (ZPEs) between the FCH ($\varepsilon_0^{\text{FCH}}$) and GS ($\varepsilon_0^{\text{GS}}$) states, i.e.,

$$\Delta\varepsilon_0 = \varepsilon_0^{\text{FCH}} - \varepsilon_0^{\text{GS}}. \quad (4)$$

Stick spectra were convoluted by a Lorentzian line shape and slightly different HWHMs were set to better compare with different experiments [4,40,41]: 0.03 eV for indole, 3-methylindole, and 3-formylindole and 0.05 eV for others. For each molecule with multiple nitrogen atoms, its total spectrum was calculated simply by summing the individual atom-specific contributions. In addition, major stick transitions were analyzed, where thresholds for Franck-Condon factors (FCFs) $F \geqslant 0.02$ (indole, 3-methylindole, and 3-formylindole) and $F \geqslant 0.04$ (the rest) were used to filter out weak transitions. All major assignments of each individual molecule are provided in the Supplemental Material [51] (see Figs. S1–S17 therein, emphasized with colored vertical lines). For those with available experiments (indole [40], 3-methylindole [41], 3-formylindole [40], and adenine [4]), theoretical spectra were shifted by 0.25, 0.32, 0.22, and 0.44 eV, respectively, to better compare with the experiments. No *ad hoc* shift was applied for the rest. All raw numerical data of the theoretical spectra and Cartesian coordinates of all optimized geometries are available online [52].

III. RESULTS

A. Statistics on vertical and adiabatic IPs

Table I lists the theoretical vertical and adiabatic N 1s IPs for the 33 nitrogen sites of all 17 molecules. The data are further visualized in Fig. 2, where a clear separation of 15 amine and 18 imine nitrogen atoms is vividly illustrated. All theoretical vertical IPs are in a range of 403.3–407.2 eV spanning 3.9 eV. All amine N atoms are located at higher-energy (405.5–407.2 eV) and all imine N atoms are at lower-energy (403.3–405.2 eV) regions, with a gap of approximately 0.3 eV. Comparisons are made with available experiments [4,40,41] for selected molecules and the deviations of vertical IPs are from -0.4 to $+0.1$ eV. The deviations are consistent with those found in ΔKohn-Sham calculations of other molecules [33,53–56].

We also compare the adiabatic IPs with experiments [4,40,41], which show larger deviations (from -0.8 to -0.2 eV) than vertical calculations. Adiabatic IPs cover a range of 3.7 eV (403.1–406.8 eV), with the amine and imine N atoms being in the higher-energy (405.1–406.8 eV) and

TABLE I. Vertical and adiabatic ionization potentials I^{vert} and I^{ad}, respectively, 0-0 transition energies E_{00}^{00}, and $\Delta\varepsilon_0$ of 17 molecules simulated by B3LYP. Here N* denotes the core-ionized N atom; the N indexes are defined in Fig. 1(a); and amine (bold) and imine (lightface) N atoms are distinguished in different fonts. Deviations from experiments are included in parentheses. All values are in eV.

Molecule	N*	Expt.	I^{vert}	I^{ad}	E_{00}^{DR}	$\Delta\varepsilon_0$
indole	**N**	405.82[a]	405.68 (−0.14)	405.29 (−0.53)	404.89	−0.02
3-methylindole	**N**	405.70[b]	405.46 (−0.24)	405.11 (−0.59)	405.01	−0.01
3-formylindole	**N**	406.36[a]	406.24 (−0.12)	405.90 (−0.46)	405.86	−0.05
adenine	N1	404.4[c]	403.91 (−0.09)	403.62 (−0.78)	403.97	+0.04
adenine	N2	404.4[c]	404.04 (−0.40)	403.74 (−0.66)	403.33	+0.04
adenine	N3	404.4[c]	404.45 (+0.05)	404.19 (−0.21)	404.22	+0.03
adenine	**N4**	406.7[c]	406.37 (−0.33)	405.87 (−0.83)	405.79	−0.08
adenine	**N5**	405.7[c]	405.47 (−0.33)	405.10 (−0.60)	405.06	−0.04
purine	N1		404.48	404.27	404.29	+0.01
purine	N2		404.83	404.60	404.60	+0.00
purine	N3		404.73	404.49	404.51	+0.02
purine	**N4**		406.75	406.28	406.18	−0.10
benzimidazole	**N1**		406.17	405.76	405.68	−0.08
benzimidazole	N2		404.00	403.78	403.82	+0.04
4-azaindole	**N1**		405.98	405.68	405.65	−0.03
4-azaindole	N2		403.81	403.63	403.67	+0.04
5-azaindole	**N1**		406.07	405.77	405.33	−0.04
5-azaindole	N2		403.76	403.57	403.62	+0.05
6-azaindole	**N1**		406.07	405.77	405.73	−0.04
6-azaindole	N2		403.77	403.60	403.64	+0.04
7-azaindole	**N1**		405.82	405.48	405.44	−0.04
7-azaindole	N2		404.10	403.92	403.95	+0.03
1H-isoindole	N		403.88	403.71	403.75	+0.04
2H-indole	N		403.31	403.13	403.18	+0.05
3H-indole	N		404.52	404.36	404.36	+0.00
2H-isoindole	**N**		406.07	405.80	404.76	−0.04
indolizine	**N**		406.18	405.92	404.99	−0.03
7-azaindazole	**N1**		406.43	405.97	405.90	−0.07
7-azaindazole	N2		405.19	405.03	405.03	+0.00
7-azaindazole	N3		404.51	404.31	404.33	+0.02
pyrazolo[1,5-a]pyrimidine	**N1**		407.15	406.75	406.68	−0.07
pyrazolo[1,5-a]pyrimidine	N2		404.66	404.45	404.66	+0.01
pyrazolo[1,5-a]pyrimidine	N3		404.80	404.60	404.62	+0.02

[a]From Ref. [40].
[b]From Ref. [41].
[c]From Ref. [4].

lower-energy (403.1–405.0 eV) regions, respectively. The two regions become very close with a gap of only 0.1 eV.

B. Statistics on ΔI

For each N 1s ionization, the adiabatic IP is about 0.2–0.5 eV lower than the corresponding vertical IP. By combining Eqs. (1) and (2), their difference can be computed by the total energy change in the FCH-state PES:

$$\Delta I \equiv I^{\text{vert}} - I^{\text{ad}} = E_{\text{FCH}}|_{\text{min GS}} - E_{\text{FCH}}|_{\text{min FCH}}. \quad (5)$$

Here ΔI describes the structural relaxation effect in the excited-state PES caused by the core ionization. This parameter approximately (neglecting the curvature change) provides an estimate of the displacement in PESs. From our calculations, it is interesting to find that ΔI for amine N atoms (0.3–0.5 eV) are generally larger than for imine N atoms (0.2–0.3 eV). We deduce that displacements in PESs of amine N atoms are generally larger than those of imine N atoms.

C. Statistics on global structural changes

This deduction is supported by Fig. 3, where root-mean-square distances (RMSDs) between the ground- and excited-state structures are plotted for each nitrogen ionization. It can be seen that ionizations of amine N atoms lead to larger RMSDs (0.03–0.10 Å) than those of imine N atoms (0.03–0.05 Å). This can be understood using that an N=C double bond (as in imines) can be more "rigid" than two N–C (or N–N) single bonds (as in amines) during the N 1s ionization process.

D. Statistics on local structural changes

In physical sciences, commonly used techniques to determine atomic structures include x-ray-diffraction and nuclear magnetic resonance techniques. However, these techniques primarily focus on ground-state structures. Determining the excited-state structures, which are transient in nature, is much more challenging. The excited structure that we dis-
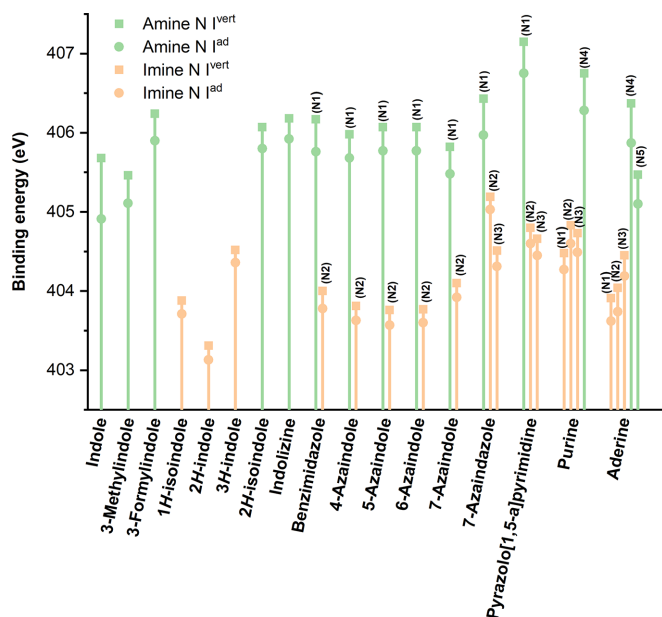
FIG. 2. Simulated vertical (I^{vert}, square) and adiabatic (I^{ad}, circle) ionization potentials of all molecules. Amine (green) and imine (orange) nitrogen atoms are distinguished by color. For the definition of each nitrogen atom see Fig. 1(a).

cussed above refers to the minimum of the core-ionized state PES. Experiments that determine such structures can be dated to decades ago. For diatomic molecules, one can utilize, for example, the Morse potential and the measured high-resolution x-ray spectrum to fit for the bond length of the excited structure (i.e., equilibrium position in the
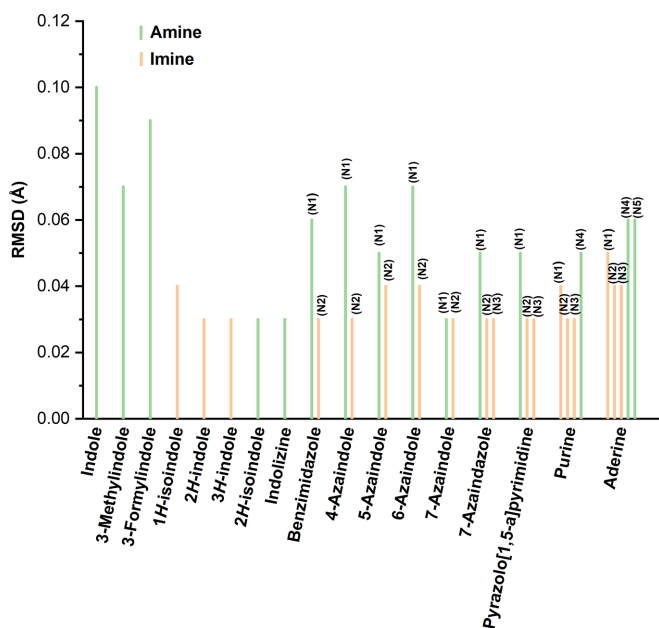

FIG. 3. RMSD values between two Cartesian coordinates, the optimized core-ionized (min FCH) and ground (min GS) states, for each N 1s ionization of each molecule. Amine (green) and imine (orange) nitrogen atoms are distinguished by color. For the definition of each nitrogen atom see Fig. 1(a).

excited-state PES) [57–59]. With the excited structure at hand, one can study whether core ionization can lead to contraction or elongation of a specific bond, aiding in the understanding of x-ray physics [60]. In recent years, advancements in science and technology, particularly ultrafast x-ray scattering with x-ray free-electron lasers [61–63] and MeV ultrafast electron diffraction [64–66], have made it possible to capture "molecular movies." These techniques offer new avenues for observing transient valence- or core-excited-state structures, which can aid in understanding ultrafast physical processes and chemical reactions.

In this study we specifically focus on examining the local geometrical changes at the N* center induced by N 1s electron ionization. Figure 4 compares the optimized structures of the ground and the core-ionized states for three selected systems: indole, purine, and adenine. It is worth noting that the left-hand side of Fig. 4 (and also Fig. 1) illustrates only one of the Kekulé structures for each molecule. A conjugated π bond system and resonant Kekulé structures average away the difference of those single and double bonds (especially in the six-membered ring) in the geometrical optimizations. We thus simply use N-C (N*-C) to denote the bond length in any range between the two atoms at min GS (min FCH). For instance, for N2 in adenine (Fig. 4), the two N2-C distances at min GS are equal (1.33 Å). The equality is broken by core ionization and the resulting N2*-C bond lengths become 1.36 and 1.33 Å, respectively. Similarly in purine, the two almost equal bond lengths (1.33 and 1.32 Å) in the ground state become 1.37 and 1.31 Å, respectively, in the core-ionized state. The structural changes induced by N1 ionization in adenine and purine are different. In purine, the two N1-C bond lengths (both 1.33 Å) stay almost unchanged after the ionization, while in adenine, the two N1-C lengths (both 1.34 Å) become 1.37 and 1.34 Å, respectively after the ionization. This is because the –NH$_2$ substitution in adenine is close to N1 and changes its local environment, which leads to different structural responses in both molecules to the 1s ionizations.

In the five-membered ring, the two N-C distances at each N site seem to agree better with the Kekulé structures, showing evident one long and one short distances. At N3 (see Fig. 4), the two N-C distances are 1.38 and 1.30 Å, respectively, in purine (1.38 and 1.31 Å in adenine), which are almost unchanged after the N 1s ionization, while evident elongation happens at N4, which corresponds to 0.07 and 0.08 Å in indole, 0.06 and 0.13 Å in purine, and 0.07 and 0.15 Å in adenine. For the three molecules, all N-H distances are slightly reduced by 0.02 Å from 1.00 Å to 0.98 Å.

Table II presents more complete data for bond lengths and angles at N* of all 17 molecules at min FCH. The N$_a$*-H bond lengths stay almost a constant value of 0.98 Å, decreasing by only 0.02 or 0.03 Å from the GS geometry. For the N*-X distances [X denotes C (mainly) or N (rarely)], amine nitrogen atoms (N$_a$*-X, 1.40–1.56 Å) always show longer values than imine nitrogen atoms (N$_i$*-X, 1.30–1.43 Å). The change of the N*-X length compared to the GS geometry is different for amine and imine N atoms: N$_a$*-X is always elongated by 0.04-0.18 Å, while N$_i$*-X can be either elongated or shortened by 0.00–0.03 Å. Concerning the bond angles ∠C–N*–X, we found a decrease for all amine N atoms (∠C–N$_a$*–X) and an increase for all imine N atoms (∠C–N$_i$*–X).
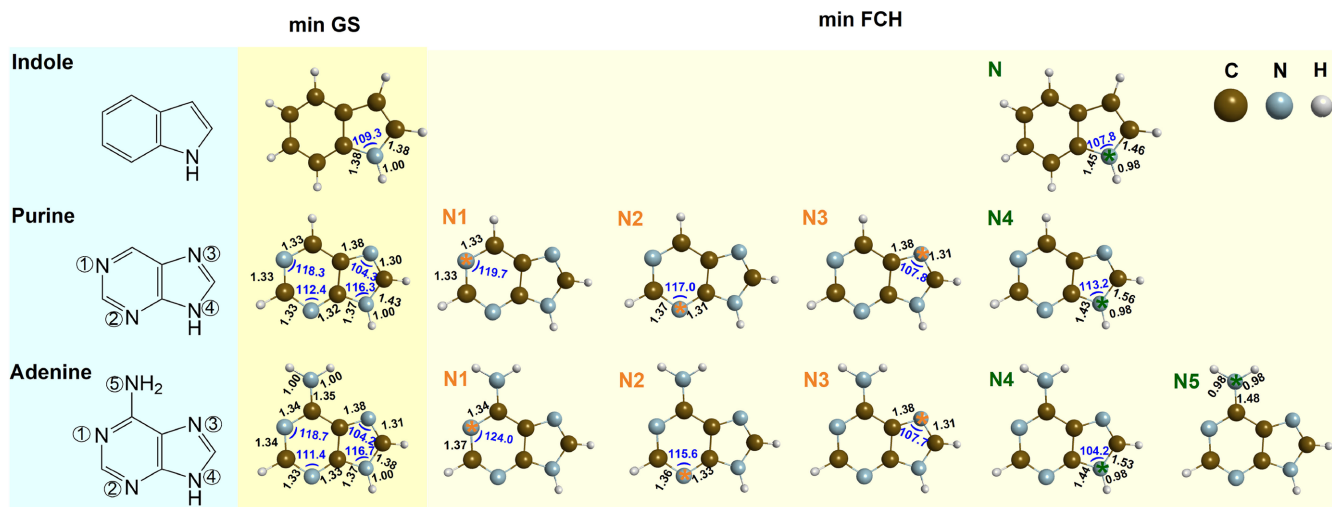
FIG. 4. Optimized geometries of three selected molecules in the ground (min GS) and core-ionized (min FCH) states, with selected bond lengths (in angstroms) and angles (in degrees) at N* are labeled. The corresponding Kekulé structures are also recaptured [from Fig. 1(a)] on the left for comparison.

E. Statistics on ZPE changes $\Delta\varepsilon_0$

To estimate the deformation of the excited-state PESs from the ground-state ones as induced by core ionization, we analyzed the $\Delta\varepsilon_0$ values (the zero-point vibrational energies in the core-ionized states with respect to the ground states) of all these 33 nonequivalent N sites. As depicted in Fig. 5, we found that amine (imine) nitrogen atoms show negative (positive) $\Delta\varepsilon_0$ values, which are about -0.1 to -0.01 (0.00 to $+0.05$) eV. Our results indicate that N $1s$ ionization leads to a distinct change in direction in the curvature of the final-state PES, which becomes flatter (steeper) for amine (imine) N atoms.

IV. DISCUSSION

A. Effects of side chain substitutions

1. –CH₃ and –CHO substitutions on indole

Figure 6(a)–6(c) depicts our computed vibrationally resolved N $1s$ XPS spectra of indole, 3-methylindole, and 3-formylindole compared to experiments [40,41]. Spectra of the three molecules exhibit significant differences. The (small, approximately 0.2 eV) redshifts and (moderate, approximately 0.4 eV) blueshifts in 3-methylindole and 3-formylindole are simply because –CH₃ and –CHO are a (weak) electron-donating and a (moderate) electron-withdrawing groups, respectively.

The experimental spectrum of indole [40] shows only an asymmetric broad peak at 405.8 eV. Our theoretical spectrum reproduces well the general profile and further identifies three characteristic peaks at 405.6, 405.8, and 405.9 eV, which arise from 0-0 and 0-1, 0-1 and 0-2, and 0-2 transitions, respectively [Fig. S1(b) in [51]]. The experimental spectrum of 3-methylindole [41] resolved two small peaks at 405.6 and 405.7 eV, with a separation of 0.1 eV. Our simulation reproduces well the fine structures, which are interpreted as 0-2 and 0-3 transitions, respectively [Fig. S2(b) in [51]]. Concerning 3-formylindole, 0-0 and 0-1 transitions contribute a feature at 406.1 eV and 0-1, 0-2, and 0-3 transitions contribute to the high-energy region at 406.2–406.4 eV [Fig. S3(b) in [51]].

Figures 6(d)–6(f) illustrate the active vibrational modes (in solid frames) of each molecule identified according to the threshold of $F \geqslant 0.02$. Two active vibrational modes are identified for each molecule [see also Figs. S1(c)–S3(c) in [51]]. For indole [Fig. 6(d)], the two modes are an N-H bending mode ν_3 (288.6 cm^{-1}) and a ring-deformation mode ν_{22} (1078.1 cm^{-1}). Both modes still can be tracked in the
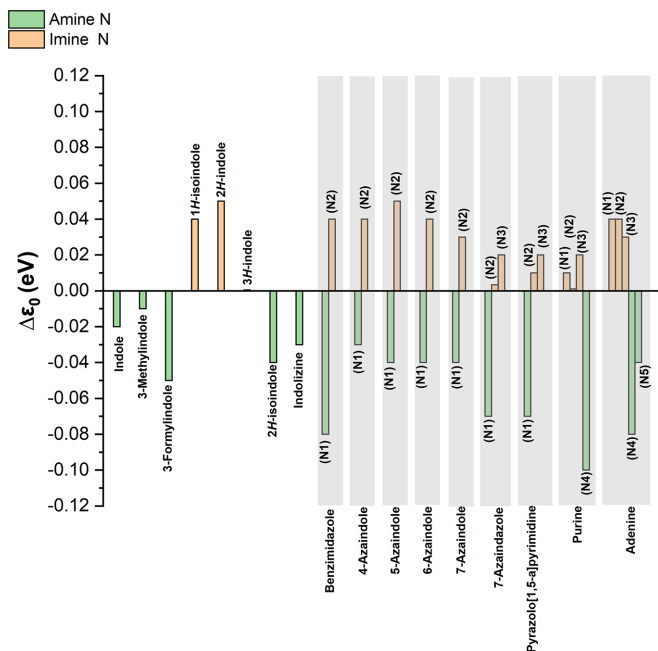

FIG. 5. The $\Delta\varepsilon_0$ values (difference between the ZPE in the final state and that in the initial state) of amine and imine nitrogen atoms in all molecules. The $\Delta\varepsilon_0$ of amine (green) and imine (orange) nitrogen atoms are distinguished by color. For the definition of each nitrogen atom see Fig. 1(a).

TABLE II. Bond lengths (in angstroms) and angles (in degrees) at each ionization center N* of each molecule (see structures and parameter definitions in Fig. 1) in its final state geometry (min FCH). For nitrogen indices, amine nitrogen N_a^* and imine nitrogen N_i^* are shown in bold and lightface fonts, respectively. Structural changes compared to the ground state (min GS) are indicated in parentheses. The RMSD values (in angstroms) of the two superimposed geometries are indicated.

Molecule	N*	N_a^*-C	N_a^*-H	N_i^*-C	∠C-N_a^*-C	∠C-N_i^*-C	RMSD
indole	**N**	1.45 (+0.07), 1.46 (+0.08)	0.98 (−0.03)		107.8 (−1.5)		0.10
3-methylindole	**N**	1.45 (+0.07), 1.47 (+0.09)	0.98 (−0.03)		107.3 (−1.8)		0.09
3-formylindole	**N**	1.45 (+0.09), 1.46 (+0.07)	0.98 (−0.03)		108.0 (−1.6)		0.07
adenine	N1			1.34 (−0.00), 1.37 (+0.03)		124.0 (+5.3)	0.05
adenine	N2			1.33 (−0.01), 1.36 (+0.03)		115.6 (+4.2)	0.04
adenine	N3			1.31 (+0.00), 1.38 (+0.00)		107.7 (+3.5)	0.04
adenine	**N4**	1.44 (+0.06), 1.53 (+0.15)	0.98 (−0.03)		104.2 (−2.6)		0.06
adenine	**N5**	1.48 (+0.13)	0.98 (−0.02)				0.06
purine	N1			1.33 (−0.00), 1.33 (−0.01)		119.7 (+1.3)	0.04
purine	N2			1.31 (−0.01), 1.37 (+0.03)		117.0 (+4.6)	0.03
purine	N3			1.31 (+0.01), 1.38 (−0.00)		107.8 (+3.5)	0.03
purine	**N4**	1.43 (+0.06), 1.56 (+0.18)	0.98 (−0.03)		103.2 (−3.1)		0.05
benzimidazole	**N1**	1.43 (+0.05), 1.55 (+0.18)	0.98 (−0.03)		103.9 (−3.0)		0.06
benzimidazole	N2			1.31 (+0.00), 1.39 (−0.00)		108.3 (+3.4)	0.03
4-azaindole	**N1**	1.44 (+0.07), 1.47 (+0.09)	0.98 (−0.03)		107.6 (−1.4)		0.07
4-azaindole	N2			1.33 (−0.01), 1.34 (+0.01)		121.4 (+5.3)	0.03
5-azaindole	**N1**	1.44 (+0.07), 1.47 (+0.08)	0.98 (−0.03)		108.0 (−1.2)		0.05
5-azaindole	N2			1.33 (+0.00), 1.34 (−0.01)		124.6 (+6.1)	0.04
6-azaindole	**N1**	1.45 (+0.07), 1.46 (+0.08)	0.98 (−0.03)		107.6 (−1.3)		0.07
6-azaindole	N2			1.33 (+0.00), 1.35 (−0.00)		124.4 (+5.8)	0.04
7-azaindole	**N1**	1.45 (+0.07), 1.46 (+0.08)	0.98 (−0.03)		107.7 (−1.3)		0.03
7-azaindole	N2			1.31 (−0.01), 1.35 (+0.02)		119.9 (+5.3)	0.03
1H-isoindole	N			1.27 (−0.01), 1.42 (−0.00)		111.0 (+4.2)	0.04
2H-indole	N			1.32 (−0.01), 1.35 (+0.02)		119.9 (+5.3)	0.03
3H-indole	N			1.32 (+0.02), 1.43 (−0.02)		109.4 (+3.8)	0.03
2H-isoindole	**N**	1.43 (+0.07), 1.43 (+0.07)	0.98 (−0.03)		110.1 (−1.9)		0.03
indolizine	**N**	1.42 (+0.05), 1.43 (+0.06), 1.47 (+0.05)			108.1 (−0.8)		0.03
7-azaindazole	**N1**	1.41 (+0.05), 1.54 (+0.18)[a]	0.98 (−0.02)		109.6 (−2.6)[c]		0.05
7-azaindazole	N2			1.33 (+0.01), 1.35 (−0.01)		108.9 (+2.2)[d]	0.03
7-azaindazole	N3			1.32 (−0.01), 1.35 (+0.02)[b]		119.5 (+5.4)	0.03
pyrazolo[1,5-a]-pyrimidine	**N1**	1.39 (+0.04), 1.47 (+0.06), 1.49 (+0.14)[a]			110.5 (−2.2)[c]		0.05
pyrazolo[1,5-a]-pyrimidine	N2			1.33 (+0.02), 1.34 (−0.01)		122.2 (+5.0)	0.03
pyrazolo[1,5-a]-pyrimidine	N3			1.35 (+0.02),[b] 1.37 (+0.03)		106.3 (+2.5)[d]	0.03

[a] N_a^*-N.
[b] N_i^*-N.
[c] ∠C-N_a^*-N.
[d] ∠C-N_i^*-N.

other two substituents, but they do not always hold as the most active modes. For instance, ν_{26} of the –CH$_3$ substituent [Fig. 6(e)] and ν_5 of the –CHO substituent [Fig. 6(f)] give FCFs of only 0.01 and 0.003, respectively (less than our threshold). Meanwhile, new active modes appear in the two substituents (compared to those in indole), including another type of N-H bending mode ν_5 (309.8 cm^{-1}) in the –CH$_3$ substituent [Fig. 6(e)] and a combined ring deformation and C=O bending mode ν_2 (156.1 cm^{-1}) in the –CHO substituent [Fig. 6(f)], both with FCFs of 0.02. The results show that side chain substitution can effectively change the active modes, providing sensitive isomer-dependent signatures.

Despite the difference in active modes, the local bond lengths and angles (at N*) of the three molecules in their FCH states are very similar (Table II). The N*-C lengths are between 1.45–1.47 Å. The angles ∠C–N*–C fall between 107.3° and 108.0°.

2. –NH$_2$ substitution on purine

Figure 7(a) displays the simulated vibrationally resolved XPS spectrum of adenine compared with the experiment [4]. The experimental spectrum features three regions I, II, and III centered at approximately 404.4, 405.7, and 406.7 eV, respectively. Our simulation agrees well with the experimental fine structure. We interpret region I for imine nitrogen atoms N1, N2, and N3 and regions II and III for amine nitrogen atoms N5 (0-1 and 0-2 transitions) and N4 (0-3 transition), respectively [Figs. S4(b)–S4(f) in [51]]. In region I, three vibronic structures are reproduced. The lowest fine structure (404.5 eV) comes from the 0-1 and 0-2 contributions of N1 and N2 [Figs. S4(b) and S4(c) in [51]] and the next two (404.6 and 404.7 eV) are, respectively, from the 0-0 and 0-1 transitions of N3 [Fig. S4(d) in [51]]. According to our threshold ($F \geqslant 0.04$), only N3 contains stick vibronic transitions with
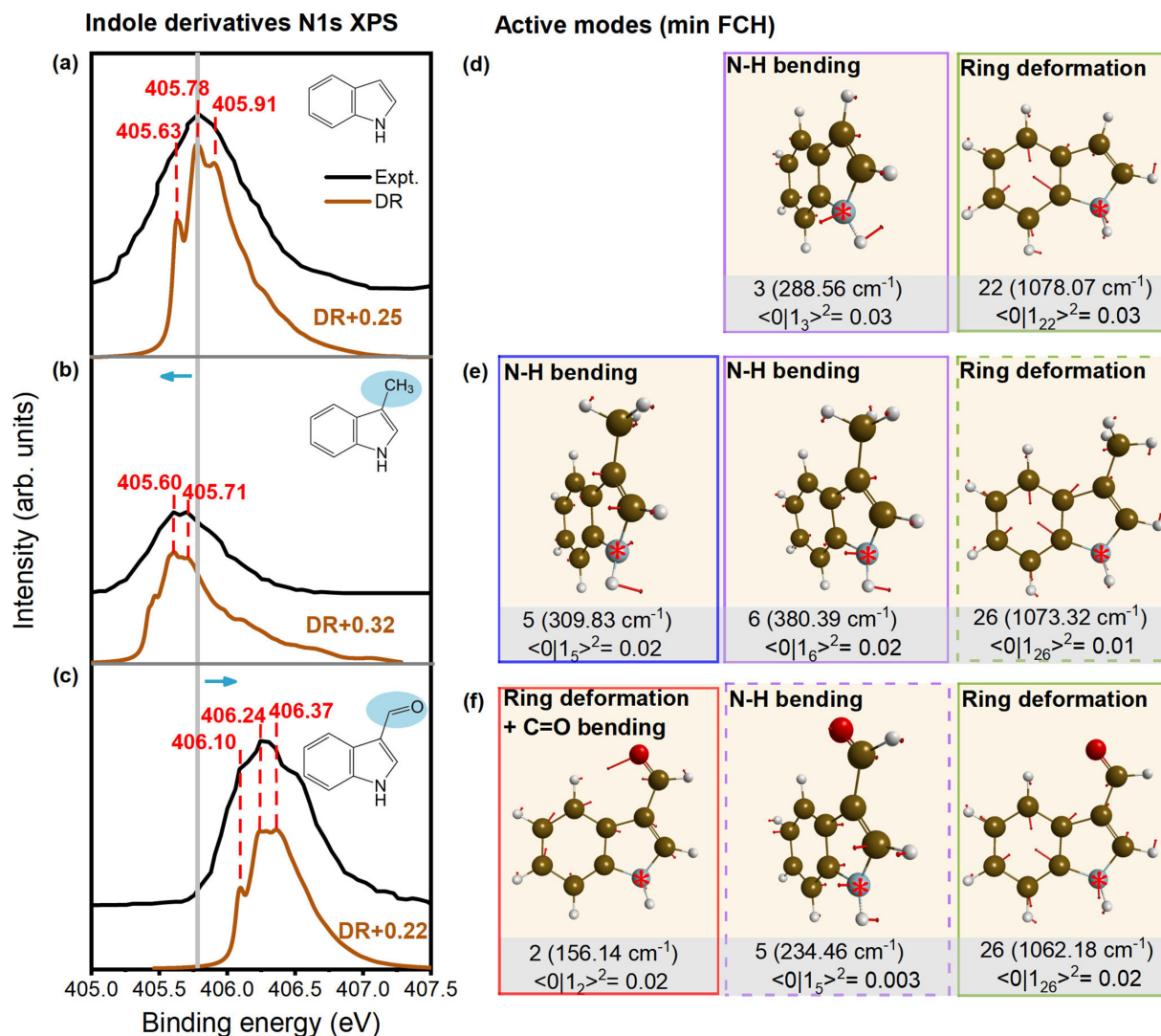
FIG. 6. Simulated vibrationally resolved N 1s XPS spectra of (a) indole, (b) 3-methylindole, and (c) 3-formylindole using the FCH-DR method. Theoretical spectra are uniformly shifted by $\delta = 0.25$, 0.32, and 0.22 eV to better compare with experiments (From Refs. [40], [41], and [40] for the three molecules, respectively). (d)–(f) Active vibration modes of the three molecules at the final-state structure (min FCH) (in solid frames, colored to show different mode types). Additionally, two less active modes are shown for comparison in dashed frames. Stars indicate N*. See the text for details.

large intensities [Fig. S4(a) in [51]]. Two active modes ν_{18} (965.4 cm^{-1}) and ν_{22} (1164.4 cm^{-1}) are identified, which are both ring deformation modes localized mainly in the five-membered ring [Fig. S4(g) in [51]].

Figure 7(b) shows the simulated spectrum of purine. Similar to adenine, imine (N1–N3) and amine (N4) nitrogen atoms contribute to the two regions I (approximately 404.6 eV) and II (approximately 406.5 eV), respectively. In region I, the three fine structures at 404.5, 404.6, and 404.7 eV come from mixed vibronic contributions by these three atoms [see also Fig. S5(a) in [51]]. In region II, the doublet fine structures with nearly equal intensities (at 406.4 and 406.5 eV) are assigned, respectively, as 0-2 and 0-3 as well as 0-3 and 0-4 transitions of N4 [Fig. S5(e) in [51]]. For each nitrogen atom, one to three active modes are identified, which are all low-frequency (330.2–1075.8 cm^{-1}) in-plane ring deformation modes [Fig. S5(f) in [51]].

Purine and adenine differ by an –NH$_2$ group. –NH$_2$ is a strong EDG, which leads to smaller (0.3–0.8 eV) vertical IP values for N1–N4 in adenine than in purine (Table I). Our calculations show that in the FCH state, the bond lengths at each of the same N* sites are similar (Table II). The N*-C lengths in each molecule deviate by 0.01-0.03 Å compared to the corresponding GS geometry. The differences in bond angles ($\angle$C–N$_a^*$–C and $\angle$C–N$_i^*$–C) are 0.1°–4.3°.

B. Effects of isomerization

Figure 8(a) depicts computed N 1s XPS spectra of six indole isomers. Above we discussed the separation of binding energies for imine and amine N atoms. Concerning the vibronic profiles, they are also distinct from each other. For each isomer, the sum of FCFs converges (a threshold of 0.99 was used) at the same $n = 7$ [Figs. S1(b) and S11(b)–
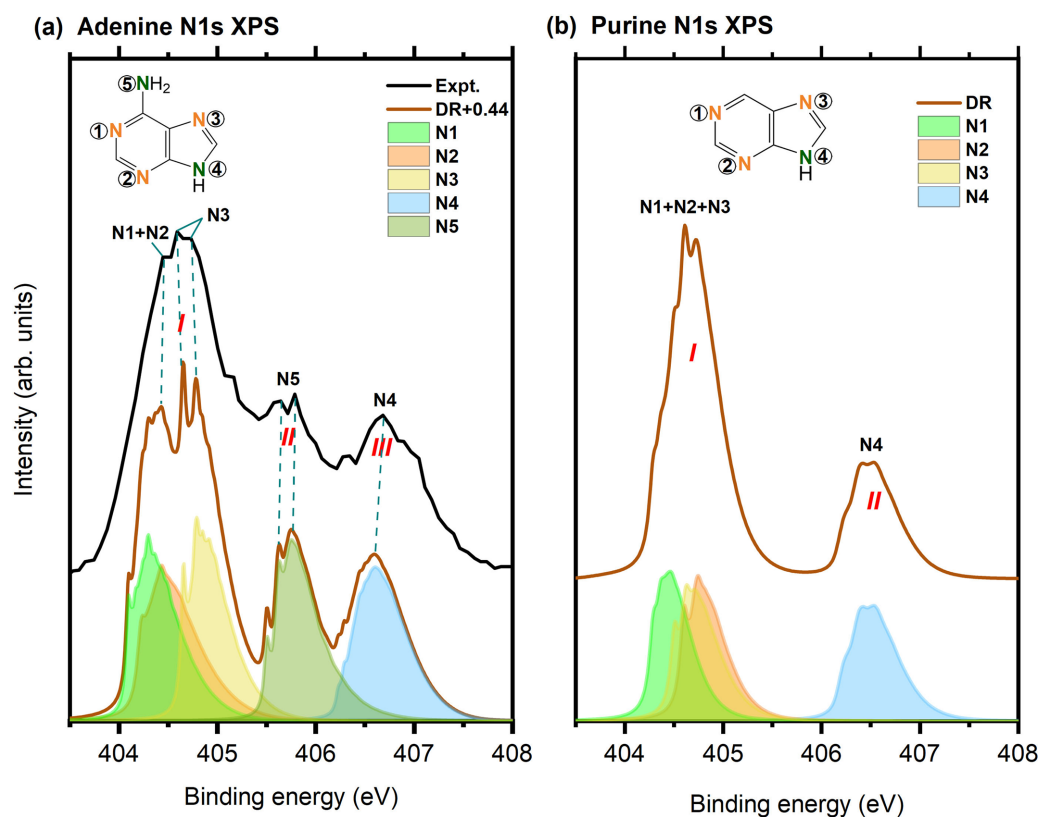
FIG. 7. Simulated vibrationally resolved N 1s XPS spectrum of (a) adenine and (b) purine. Regions are labeled by arabic letters (I for imine N and II and III for amine N) and spectral fine structures are interpreted. Shaded areas are atom-specific contributions (definitions are in each legend). In (a) the theoretical spectrum is shifted by +0.44 eV for better comparison with the experiment [4].

S15(b) in [51]]. The isomer-dependent signatures come from different modulations of each 0-n transition. For each isomer, a lower-energy 0-0 peak always exists, which is well separated from the broad peak in the higher-energy part. The broad peak is assigned to be mainly 0-1 transitions (3H-indole) or a combination of 0-1 and 0-2 transitions (indole, 1H-isoindole, 2H-indole, 2H-isoindole, and indolizine). Decomposition analyses also show that for each molecule, 0-1 transitions always make the largest contribution among all 0-n transitions. This indicates a relatively small displacement of the PESs as induced by core ionization, owing to the stiffness of the bicyclic molecules. Similar profile shapes were observed in monocyclic molecules [17,19]. Interpretation of major stick transitions of each isomer indicate that only a few (one to three) low-frequency (288.6–1507.4 cm^{-1}), nearly in-plane, ring-deformation modes are active [Figs. S1(c) and S11(c)–S15(c) in [51]].

C. Effects of mode mixing

The Duschinsky rotation matrix **J** carries information about the strengths of mode mixing. The matrix elements of each indole isomer are visualized in Fig. 8(b). Note that the range of each entry J_{ij} is from −1 to 1. The absolute value of the element $|J_{ij}|$ that is close to 0 and 1 indicates weak and strong mixing effects for the corresponding mode pairs i and j. Note that in the literature, absolute [67–69] or squared [70] values are also often used (thus giving a range from 0 to 1).

The three presentation methods are consistent, which will not influence our discussion.

As visualized in Fig. 8(b), elements with the largest absolute values are always at or near the diagonal. As we know, core ionization can lead to interchanges of the mode order. It is interesting to find that in Figs. 8(b iv)–8(b vi) (with amine N atoms) the corresponding matrix elements distribute farther off the diagonal than in Figs. 8(b i)–8(b iii) (with imine N atoms). This indicates a stronger mode mixing effect in amine N atoms than in imine N atoms. In the three molecules in Figs. 8(b iv)–8(b vi) with amine nitrogen atoms, the mode mixing happens between modes 25 and 35 (1178.5–1426.9 cm^{-1}), corresponding to nearly-in-plane ring-deformation vibrations. Analyses of the Duschinsky matrices for other molecules are provided in Figs. S18–S20 in [51], where the same conclusion for imine and amine nitrogen atoms generally applies.

D. Effects of CH ↔ N replacement

Figure 9 shows the spectra of five azaindole isomers, each with one amine nitrogen N1 and one imine nitrogen N2. The N1 is always at a fixed position, while N2 replaces the –CH group at different sites. The effects of isomerization were discussed above for indoles (Sec. IV B). Here the CH ↔ N replacement is a special kind of isomerization. Besides this effect, another theme here is to analyze the effects of N1-N2 interactions on the spectra, as discussed below (Sec. IV E).
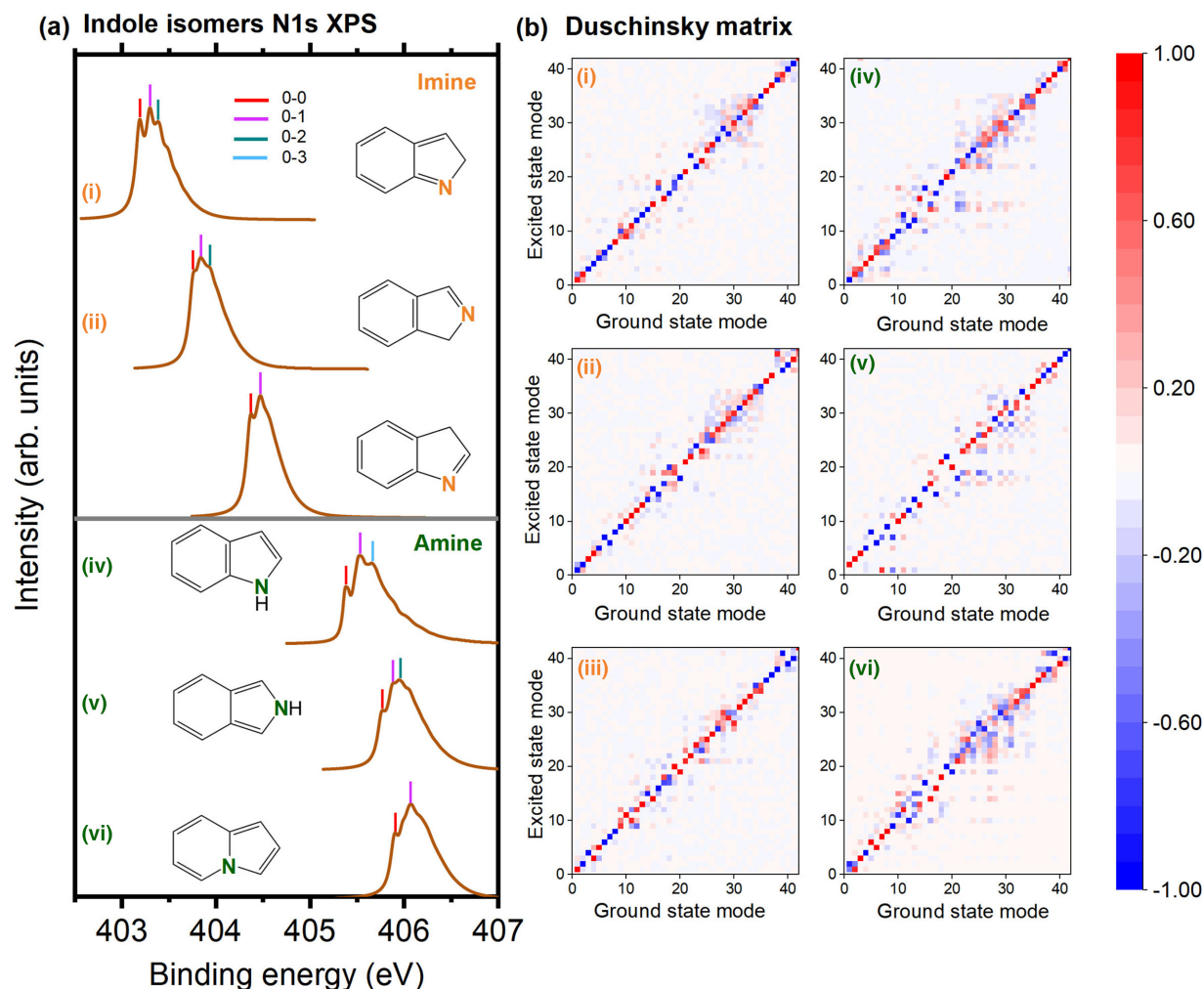
FIG. 8. (a) Comparison of simulated vibrationally resolved N 1s XPS spectra of six indole isomers: (i) 2H-indole, (ii) 1H-isoindole, (iii) 3H-indole, (iv) indole, (v) 2H-isoindole, and (vi) indolizine for (i)–(iii) an imine N atom and (iv)–(vi) an amine N atom. (b) Duschinsky matrix of each isomer. The elements are generally more "diffused" off the diagonals for amine (right) than imine (left) nitrogen atoms.

For each isomer, the two peaks from N1 and N2 are clearly separated. Responding to the position change, the N2 peak varies within an energy range of 0.3 eV between 403.8 and 404.1 eV and its fine structure shows a clear isomer-dependent signature. Each molecule is identified to have two to three active vibrational modes. They are all in-plane modes within a frequency range of 246.5–1576.2 cm^{-1} [Figs. S6(d)–S10(d) in [51]]. For m-azaindole (m = 4, 5, 6, 7), all active modes of arise from N2. Structurally, benzimidazole is the only molecule that contains both nitrogen atoms in the five-membered pyrrole ring. It is interesting to find that the active modes of benzimidazole are ring deformation vibrations localized only on the five-membered ring [Fig. S6(d) in [51]]. For other isomers, the actives are more delocalized [Figs. S7(d)–S10(d) in [51]].

In addition, the FCH-optimized structure of benzimidazole is clearly distinguished from the other four isomers (Table II). The N1*-C bond length of benzimidazole (1.55 Å) is significantly greater than that of the four m-azaindole isomers (1.45-1.47 Å). The same conclusion applies for N2: The N2*-C distance in benzimidazole is 1.39 Å, while in rest isomers, the length is 1.33–1.34 Å. Similarly, we also

find that the bond angle $\angle$C–N$_a^*$–C ($\angle$C-N$_i^*$–C) for the amine N1 (imine N2) is 107.6°–108.0° (121.4°–124.6°) for the four m-azaindole isomers, while for benzimidazole the value is much smaller: 103.9° (108.3°). Results indicate that the influence of CH $\leftrightarrow$ N replacement within a five- or a six-membered ring is evidently different.

E. Effects of N-N interactions

Although N1 is always fixed in each azaindole isomer, it is interesting to find that the energy position of the N1 peak also varies evidently by 0.4 eV (from 405.8 to 406.2 eV), a value comparable to that of N2. The corresponding spectral profiles of the N1 peaks in each molecule are also distinct from each other. The change is attributed to the structural perturbation caused by N2. This reflects the interaction between the two nitrogen atoms, which varies by molecules. The atom-atom interaction within a molecule was theoretically investigated and compared in p-, m-, and o-aminophenols with para-, meta-, and ortho-positions of the –NH$_2$ and –OH groups by core excitations [71]. Here the two nitrogen atoms are connected via the delocalized π orbital since alternating single and
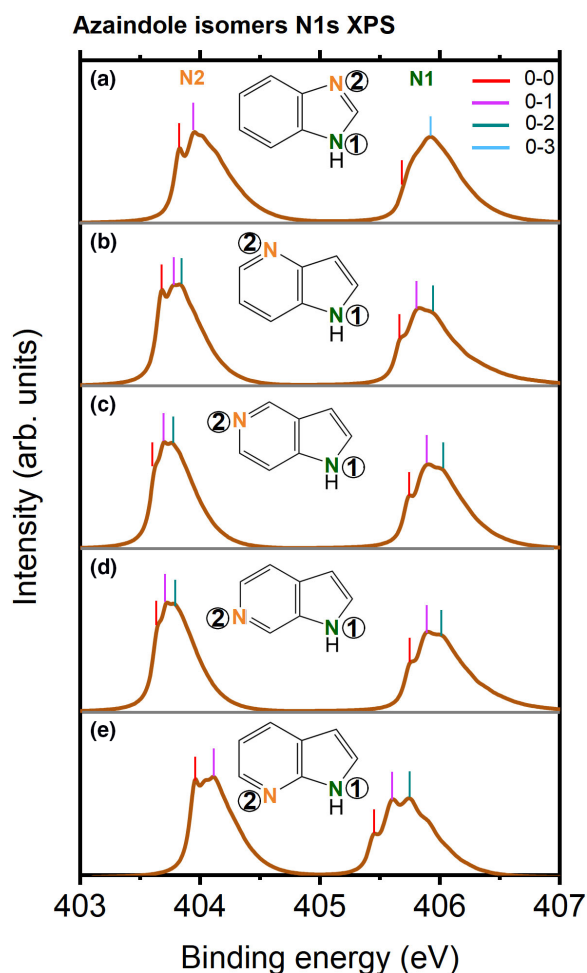
FIG. 9. Comparison of simulated vibrationally resolved N 1*s* XPS spectra of five azaindole isomers. In each isomer, the amine nitrogen N1 is always fixed while the imine nitrogen N2 is altered: (a) benzimidazole, (b) 4-azaindole, (c) 5-azaindole, (d) 6-azaindole, and (e) 7-azaindole. Peaks are interpreted.

double bonds create a conjugated π bond system across multiple atoms.

From another point of view, the interaction between the two nitrogen atoms can be read by comparing the spectra of 5- and 6-azaindole [Figs. 8(b iii) and 8(b iv)]. In these two molecules, N1 and N2 have the largest spatial separation and consequently the weakest interaction (similar to the role of p-aminophenol among all aminophenols [71]); the resulting spectra (both the energies and profiles) are close to each other.

F. Limitation of the FCH DR approach

Another extreme in terms of the nitrogen-nitrogen interaction is when two N atoms directly bond to each other (the strongest interaction). Within a five-membered ring, this results in another isomer, 1*H*-indazole. We tried to simulate its spectrum with the same protocol, but the optimized FCH-state geometry deviates much from the GS one (a much larger N*-N distance). As a result, the consequent Franck-Condon calculations failed to converge. The existence of the N*-N motif seems to indicate a difficult case for FCH DFT optimizations, though it does not always fail.

Critically, the same problem was also encountered in 1,2,3-triazine, 1,2,3,4- and 1,2,3,5-tetrazine [17], and N_5H [33] (these molecules were abandoned in the final publications due to the failure). The essential problem lies in locating the minima of the core-ionized isomer (min FCH), while a single-point (vertical) FCH calculation at min GS usually works. This could be the limitation of the FCH DR method when there is a tendency to significantly elongate the bond length at the ionization center. An approximate solution is to use the linear coupling model (LCM) [72] instead of the Duschinsky rotation, since the LCM calculation can avoid locating the structure of min FCH [6].

In general, the accuracy of vibrationally resolved electronic spectra depends on the electronic structure method and the vibronic coupling model. The FCH DR method has been shown to provide accurate spectra for many common cyclic molecules [17,19,20]. For difficult cases, high-level electronic structure methods such as the multiconfigurational methods [73] can be tried and Hertzberg-Teller effects beyond the Franck-Condon approximation can be included [74]. For smaller systems or limited degrees of freedom, a time-dependent framework can be used to scan the PESs and consider anharmonic effects [75]. Another possible reason with the Franck-Condon analysis for the difficult cases is the use of the rectilinear coordinates, which could overestimate the displacements between the PESs of the two states involved in the electronic transition process. This effect is particularly pronounced in systems where there are significant geometric changes before and after the electronic transition. To improve accuracy, one possible approach is to employ curvilinear coordinates, as proposed by Reimers [76].

G. Timing

The computational cost was examined on a computing node containing four dual-core Intel Xeon processors E5-2667 v3 (3.20 GHz). The calculation generally includes the electronic structure and spectral calculation steps [19]. The electronic structure step contains mainly geometrical optimization and vibrational frequency calculations of both the ground and core-ionized states. For these molecules with 13–19 atoms, it normally takes 50–60 h for this step if we run the jobs sequentially (i.e., the ground state followed by core-excited state and optimization followed by frequency). The excited-state calculation is usually about 20% longer than the corresponding ground-state calculation. Concerning the spectral calculation step, with the OPENMP parallelization [65] implemented in the DYNAVIB package [48], it takes only a few minutes.

V. CONCLUSION

In summary, we have computed the vibrationally resolved XPS spectra for a family of 17 indole-based bicyclic molecules at the DFT level to investigate the structure-spectroscopy relation and to yield general vibronic coupling rules of this family. Structural variations cover common side chain substitutions, isomerizations, and CH $\leftrightarrow$ N replacements. Vibronic coupling was considered by the Franck-Condon approach with the inclusion of the Duschinsky rotation effect. Good agreement with available experi-

ments was achieved for both absolute binding energies and fine structures, which validated reliable analyses. Binding energies, major vibronic transitions, corresponding active vibrational modes, and core-ionization-induced local geometrical changes at N^* were analyzed for each molecule. Our study in this work provides reliable and complete spectral data for future experimental and theoretical references and machine learning studies, and the rules summarized are insightful for a better understanding of the basic physical concept of vibronic coupling in the XPS process.

In this work, not only was the clear separation of binding energies of amine (larger) and imine (smaller) nitrogen atoms, as reported in several early works [33,35], confirmed, but more importantly it is interesting to find the same separation rules also for vibronic properties. (1) The $\Delta\varepsilon_0$ value (zero-point vibrational energy of the FCH state compared to the ground state) is positive for all amine N atoms and negative for all imine N atoms. This indicates that the change in the curvature of the excited-state PES induced by core ionization depends on the local structure at the N^* center, which can be steeper (for amine N atoms) or smoother (for imine N atoms). (2) Amine nitrogen atoms exhibit stronger mode mixing effects than imine N atoms. (3) The $1s$ ionization of amine N atoms always leads to an elongation of N^*-C and shortening of the N^*-H bond lengths, while for imine N atoms, the structural changes in N^*-C are much smaller and both elongation and shortening exist. (4) Ionization of amine (imine) nitrogen atoms always leads to a decrease (increase) in the bond angle $\angle$C–N^*–C. (5) The $1s$ ionization in an amine N atom leads to a larger global structural change (indicated by the RMSD between min GS and min FCH) than in an imine N atom.

ACKNOWLEDGMENTS

[1] U. Hergenhahn, Vibrational structure in inner shell photoionization of molecules, J. Phys. B **37**, R89 (2004).

[2] M. Mendolicchio, A. Baiardi, G. Fronzoni, M. Stener, C. Grazioli, M. De Simone, and V. Barone, Theory meets experiment for unravelling the C1s X-ray photoelectron spectra of pyridine, 2-fluoropyridine, and 2,6-difluoropyridine, J. Chem. Phys. **151**, 124105 (2019).

[3] O. Plekan, C. Grazioli, M. Coreno, M. Di Fraia, K. C. Prince, R. Richter, and A. Ponzi, Investigation of quinoline derivatives by photoemission spectroscopy and theoretical calculations, Chem. Phys. **562**, 111657 (2022).

[4] O. Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K. Prince, A. Trofimov, E. Gromov, I. Zaytseva, and J. Schirmer, A theoretical and experimental study of the near edge X-ray absorption fine structure (NEXAFS) and X-ray photoelectron spectra (XPS) of nucleobases: Thymine and adenine, Chem. Phys. **347**, 360 (2008) See, in particular, Fig. 11 therein, showing the experimental N $1s$ XPS spectrum of adenine in the gas phase, which has been recaptured in our Fig. 5.

[5] V. Myrseth, K. J. Børve, K. Wiesner, M. Bässler, S. Svensson, and L. J. Sæthre, Vibrational structure and vibronic coupling in the carbon 1s photoelectron spectra of benzene and deuterobenzene, Phys. Chem. Chem. Phys. **4**, 5937 (2002).

[6] G. Fronzoni, O. Baseggio, M. Stener, W. Hua, G. Tian, Y. Luo, B. Apicella, M. Alfé, M. de Simone, A. Kivimäki, and M. Coreno, Vibrationally resolved high-resolution NEXAFS and XPS spectra of phenanthrene and coronene, J. Chem. Phys. **141**, 044313 (2014).

[7] R. Sankari, M. Ehara, H. Nakatsuji, Y. Senba, K. Hosokawa, H. Yoshida, A. De Fanis, Y. Tamenori, S. Aksela, and K. Ueda, Vibrationally resolved O 1s photoelectron spectrum of water, Chem. Phys. Lett. **380**, 647 (2003).

[8] K. Siegbahn, Electron spectroscopy for atoms, molecules, and condensed matter, Rev. Mod. Phys. **54**, 709 (1982).

[9] K. Siegbahn, C. Nordling, A. Fahlman, R. Nordberg, K. Hamrin, J. Hedman, G. Johansson, T. Bergmark, S. Karlsson, I. Lindgren *et al.*, *ESCA: Atomic, Molecular and Solid State Structure Studies by Means of Electron Spectroscopy* (Almqvist & Wiksell, Uppsala, 1967).

[10] K. Siegbahn, D. A. Allison, and J. H. Allison, *ESCA-photoelectron Spectroscopy* (CRC, Cleveland, 1974).

[11] J. R. Rumble, Jr., D. M. Bickham, and C. J. Powell, The NIST x-ray photoelectron spectroscopy database, Surf. Interface Anal. **19**, 241 (1992).

[12] C. D. Wagner, *Handbook of X-Ray Photoelectron Spectroscopy: A Reference Book of Standard Data for Use in X-Ray Photoelectron Spectroscopy* (Perkin-Elmer, Waltham, 1979).

[13] W. Jolly, K. Bomben, and C. Eyermann, Core-electron binding energies for gaseous atoms and molecules, At. Data Nucl. Data Tables **31**, 433 (1984).

[14] The International XPS Database, https://xpsdatabase.com/.

[15] http://www.sasj.jp/.

[16] http://www.lasurface.com/xps/index.php/.

[17] M. Wei, X. Cheng, L. Zhang, J.-R. Zhang, S.-Y. Wang, G. Ge, G. Tian, and W. Hua, Vibronic fine structure in the nitrogen $1s$ photoelectron spectra of molecules from Franck-Condon simulations: Azines, Phys. Rev. A **106**, 022811 (2022).

[18] G. Greczynski and L. Hultman, X-ray photoelectron spectroscopy: Towards reliable binding energy referencing, Prog. Mater. Sci. **107**, 100591 (2020).

[19] W. Hua, G. Tian, and Y. Luo, Theoretical assessment of vibrationally resolved C1s X-ray photoelectron spectra of simple cyclic molecules, Phys. Chem. Chem. Phys. **22**, 20014 (2020).

[20] X. Cheng, M. Wei, G. Tian, Y. Luo, and W. Hua, Vibrationally-resolved x-ray photoelectron spectra of six polycyclic aromatic hydrocarbons from first-principles simulations, J. Phys. Chem. A **126**, 5582 (2022).

[21] J. A. Joule, *Heterocyclic Chemistry* (CRC, Boca Raton, 2020).

[22] Y. Damour, M. Véril, F. Kossoski, M. Caffarel, D. Jacquemin, A. Scemama, and P.-F. Loos, Accurate full configuration interaction correlation energy estimates for five- and six-membered rings, J. Chem. Phys. **155**, 134104 (2021); G. da Silva and J. W.

Bozzelli, Retro-[3+2]-cycloaddition reactions in the decomposition of five-membered nitrogen-containing heterocycles, J. Org. Chem. **73**, 1343 (2008); K. E. Gutowski, R. D. Rogers, and D. A. Dixon, Accurate thermochemical properties for energetic materials applications. I. Heats of formation of nitrogen-containing heterocycles and energetic precursor molecules from electronic structure theory, J. Phys. Chem. A **110**, 11890 (2006); S. P. Verevkin, V. N. Emel'yanenko, R. Notario, M. V. Roux, J. S. Chickos, and J. F. Liebman, Rediscovering the wheel. Thermochemical analysis of energetics of the aromatic diazines, *ibid.* **3**, 3454 (2012).

[23] N. Kerru, L. Gummidi, S. Maddila, K. K. Gangu, and S. B. Jonnalagadda, A review on recent advances in nitrogen-containing molecules and their biological applications, Molecules **25**, 1909 (2020).

[24] Y. Zheng, X. Qi, S. Chen, S. Song, Y. Zhang, K. Wang, and Q. Zhang, Self-assembly of nitrogen-rich heterocyclic compounds with oxidants for the development of high-energy materials, ACS Appl. Mater. Interfaces **13**, 28390 (2021).

[25] M. R. Castro, P. Schellenberg, M. Belsley, A. C. Fonseca, S. S. Fernandes, and M. M. Raposo, Design, synthesis and evaluation of redox, second order nonlinear optical properties and theoretical DFT studies of novel bithiophene azo dyes functionalized with thiadiazole acceptor groups, Dyes Pigm. **95**, 392 (2012).

[26] A. Mermer, T. Keles, and Y. Sirin, Recent studies of nitrogen containing heterocyclic compounds as novel antiviral agents: A review, Bioorg. Chem. **114**, 105076 (2021).

[27] P. Bolognesi, P. O'Keeffe, Y. Ovcharenko, M. Coreno, L. Avaldi, V. Feyer, O. Plekan, K. C. Prince, W. Zhang, and V. Carravetta, Pyrimidine and halogenated pyrimidines near edge x-ray absorption fine structure spectra at C and N K-edges: Experiment and theory, J. Chem. Phys. **133**, 034302 (2010).

[28] R. Sundberg, *The Chemistry of Indoles* (Elsevier, Amsterdam, 2012), Vol. 18.

[29] G. W. Gribble, *Indole Ring Synthesis: From Natural Products to Drug Discovery* (Wiley, New York, 2016).

[30] Y. Jia, X. Wen, Y. Gong, and X. Wang, Current scenario of indole derivatives with potential anti-drug-resistant cancer activity, Eur. J. Med. Chem. **200**, 112359 (2020).

[31] A. Dorababu, Indole – a promising pharmacophore in recent antiviral drug discovery, RSC Med. Chem. **11**, 1335 (2020).

[32] Y. Han, W. Dong, Q. Guo, X. Li, and L. Huang, The importance of indole and azaindole scaffold in the development of antitumor agents, Eur. J. Med. Chem. **203**, 112506 (2020).

[33] X. Du, S.-Y. Wang, M. Wei, J.-R. Zhang, G. Ge, and W. Hua, A theoretical library of N1s core binding energies of polynitrogen molecules and ions in the gas phase, Phys. Chem. Chem. Phys. **24**, 8196 (2022).

[34] W. Hua, H. Yamane, B. Gao, J. Jiang, S. Li, H. S. Kato, M. Kawai, T. Hatsui, Y. Luo, N. Kosugi, and H. Ågren, Systematic study of soft X-ray spectra of poly(Dg)poly(Dc) and poly(Da)poly(Dt) DNA duplexes, J. Phys. Chem. B **114**, 7016 (2010).

[35] J.-R. Zhang, Y. Ma, S.-Y. Wang, J. Ding, B. Gao, E. Kan, and W. Hua, Accurate K-edge X-ray photoelectron and absorption spectra of g-C$_3$N$_4$ nanosheets by first-principles simulations and reinterpretations, Phys. Chem. Chem. Phys. **21**, 22819 (2019).

[36] T. Liljefors, P. Krogsgaard-Larsen, and U. Madsen, *Textbook of Drug Design and Discovery* (CRC, Boca Raton, 2002);

H. Ebrahimi, J. Hadi, and H. Al-Ansari, A new series of Schiff bases derived from sulfa drugs and indole-3-carboxaldehyde: Synthesis, characterization, spectral and DFT computational studies, J. Mol. Struct. **1039**, 37 (2013).

[37] N. A. Meanwell, M. R. Krystal, B. Nowicka-Sans, D. R. Langley, D. A. Conlon, M. D. Eastgate, D. M. Grasela, P. Timmins, T. Wang, and J. F. Kadow, Inhibitors of HIV-1 attachment: The discovery and development of temsavir and its prodrug fostemsavir, J. Med. Chem. **61**, 62 (2018).

[38] H.-Y. Lee, A.-C. Tsai, M.-C. Chen, P.-J. Shen, Y.-C. Cheng, C.-C. Kuo, S.-L. Pan, Y.-M. Liu, J.-F. Liu, T.-K. Yeh, J.-C. Wang, C.-Y. Chang, J.-Y. Chang, and J.-P. Liou, Azaindolyl-sulfonamides, with a more selective inhibitory effect on histone deacetylase 6 activity, exhibit antitumor activity in colorectal cancer HCT116 cells, J. Med. Chem. **57**, 4009 (2014).

[39] S. R. Walker, E. J. Carter, B. C. Huff, and J. C. Morris, Variolins and related alkaloids, Chem. Rev. **109**, 3080 (2009).

[40] 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 *et al.*, Experimental and theoretical photoemission study of indole and its derivatives in the gas phase, J. Phys. Chem. A **124**, 4115 (2020) See, in particular, Fig. 4 therein, showing the experimental N 1s XPS spectra of indole and 3-formylindole in the gas phase, which have been recaptured in our Fig. 4.

[41] W. Zhang, V. Carravetta, O. Plekan, V. Feyer, R. Richter, M. Coreno, and K. C. Prince, Electronic structure of aromatic amino acids studied by soft x-ray spectroscopy, J. Chem. Phys. **131**, 035103 (2009) See, in particular, Fig. 4 therein, showing the experimental N 1s XPS spectrum of 3-methylindole in the gas phase, which has been recaptured in our Fig. 4.

[42] L. He, S. Malerz, F. Trinter, S. Trippel, L. Tomaník, M. Belina, P. Slavíček, B. Winter, and J. Küpper, Specific versus non-specific solvent interactions of a biomolecule in water, arXiv:2205.08217.

[43] F. Wang, Q. Zhu, and E. P. Ivanova, Inner-shell chemical shift of DNA/RNA bases and inheritance from their parent purine and pyrimidine, J. Synchrotron Radiat. **15**, 624 (2008).

[44] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, General atomic and molecular electronic structure system, J. Comput. Chem. **14**, 1347 (1993).

[45] M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, Amsterdam, 2005), pp. 1167–1189.

[46] A. D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, Phys. Rev. A **38**, 3098 (1988); A new mixing of Hartree–Fock and local density-functional theories, J. Chem. Phys. **98**, 1372 (1993); C. Lee, W. Yang, and R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, Phys. Rev. B **37**, 785 (1988).

[47] F. Duschinsky, The importance of the electron spectrum in multi atomic molecules. Concerning the Franck-Condon principle, Acta Physicochim. URSS **7**, 551 (1937).

[48] G. Tian, S. Duan, W. Hua, and Y. Luo, DynaVib, version 1.0, Royal Institute of Technology: Sweden, 2012.

[49] T. H. Dunning, Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, J. Chem. Phys. **90**, 1007 (1989).

[50] R. A. Kendall, T. H. Dunning, and R. J. Harrison, Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions, J. Chem. Phys. **96**, 6796 (1992).

[51] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevA.108.022816 for the spectrum of each individual molecule with analysis and visualization of the Duschinsky matrices.

[52] http://jszy.njust.edu.cn/lxy/hwj_en/list.psp.

[53] P. S. Bagus, C. Sousa, and F. Illas, Consequences of electron correlation for XPS binding energies: Representative case for C(1s) and O(1s) XPS of CO, J. Chem. Phys. **145**, 144303 (2016).

[54] N. P. Bellafont, F. Illas, and P. S. Bagus, Validation of Koopmans' theorem for density functional theory binding energies, Phys. Chem. Chem. Phys. **17**, 4015 (2015).

[55] N. P. Bellafont, P. S. Bagus, and F. Illas, Prediction of core level binding energies in density functional theory: Rigorous definition of initial and final state contributions and implications on the physical meaning of Kohn-Sham energies, J. Chem. Phys. **142**, 214102 (2015).

[56] N. Pueyo Bellafont, F. Viñes, and F. Illas, Performance of the TPSS functional on predicting core level binding energies of main group elements containing molecules: A good choice for molecules adsorbed on metal surfaces, J. Chem. Theory Comput. **12**, 324 (2016).

[57] M. Neeb, J.-E. Rubensson, M. Biermann, and W. Eberhardt, Coherent excitation of vibrational wave functions observed in core hole decay spectra of O_2, N_2 and CO, J. Electron Spectrosc. Relat. Phenom. **67**, 261 (1994).

[58] R. Püttner, I. Dominguez, T. J. Morgan, C. Cisneros, R. F. Fink, E. Rotenberg, T. Warwick, M. Domke, G. Kaindl, and A. S. Schlachter, Vibrationally resolved O 1s core-excitation spectra of CO and NO, Phys. Rev. A **59**, 3415 (1999).

[59] C. T. Chen, Y. Ma, and F. Sette, *K*-shell photoabsorption of the N_2 molecule, Phys. Rev. A **40**, 6737 (1989).

[60] T. Karlsen and K. J. Børve, Accurate and approximate calculations of Franck–Condon intensities in the carbon 1s photoelectron spectrum of methane, J. Chem. Phys. **112**, 7979 (2000).

[61] M. P. Minitti, J. M. Budarz, A. Kirrander, J. S. Robinson, D. Ratner, T. J. Lane, D. Zhu, J. M. Glownia, M. Kozina, H. T. Lemke, M. Sikorski, Y. Feng, S. Nelson, K. Saita, B. Stankus, T. Northey, J. B. Hastings, and P. M. Weber, Imaging Molecular Motion: Femtosecond X-Ray Scattering of an Electrocyclic Chemical Reaction, Phys. Rev. Lett. **114**, 255501 (2015).

[62] H. Yong, X. Xu, J. M. Ruddock, B. Stankus, A. M. Carrascosa, N. Zotev, D. Bellshaw, W. Du, N. Goff, Y. Chang, S. Boutet, S. Carbajo, J. E. Koglin, M. Liang, J. S. Robinson, A. Kirrander, M. P. Minitti, and P. M. Weber, Ultrafast X-ray scattering offers a structural view of excited-state charge transfer, Proc. Natl. Acad. Sci. USA **118**, e2021714118 (2021).

[63] M. R. Ware, J. M. Glownia, N. Al-Sayyad, J. T. O'Neal, and P. H. Bucksbaum, Characterizing dissociative motion in time-resolved x-ray scattering from gas-phase diatomic molecules, Phys. Rev. A **100**, 033413 (2019).

[64] T. J. A. Wolf, D. M. Sanchez, J. Yang, R. M. Parrish, J. P. F. Nunes, M. Centurion, R. Coffee, J. P. Cryan, M. Gühr,

K. Hegazy, A. Kirrander, R. K. Li, J. Ruddock, X. Shen, T. Vecchione, S. P. Weathersby, P. M. Weber, K. Wilkin, H. Yong, Q. Zheng *et al.*, The photochemical ring-opening of 1,3-cyclohexadiene imaged by ultrafast electron diffraction, Nat. Chem. **11**, 504 (2019).

[65] J. Yang, X. Zhu, T. J. A. Wolf, Z. Li, J. P. F. Nunes, R. Coffee, J. P. Cryan, M. Gühr, K. Hegazy, T. F. Heinz, K. Jobe, R. Li, X. Shen, T. Veccione, S. Weathersby, K. J. Wilkin, C. Yoneda, Q. Zheng, T. J. Martinez, M. Centurion *et al.*, Imaging CF_3I conical intersection and photodissociation dynamics with ultrafast electron diffraction, Science **361**, 64 (2018).

[66] J. Yang, X. Zhu, J. P. F. Nunes, J. K. Yu, R. M. Parrish, T. J. A. Wolf, M. Centurion, M. Gühr, R. Li, Y. Liu, B. Moore, M. Niebuhr, S. Park, X. Shen, S. Weathersby, T. Weinacht, T. J. Martinez, and X. Wang, Simultaneous observation of nuclear and electronic dynamics by ultrafast electron diffraction, Science **368**, 885 (2020).

[67] M.-L. Hebestreit, M. Schneider, H. Lartian, V. Betz, M. Heinrich, M. Lindic, M. Y. Choi, and M. Schmitt, Structures, dipole moments and excited state lifetime of isolated 4-cyanoindole in its ground and lowest electronically excited singlet states, Phys. Chem. Chem. Phys. **21**, 14766 (2019).

[68] C. Henrichs, M. Reineke, M.-L. Hebestreit, and M. Schmitt, Excited state structure of isolated 4-cyanoindole from a combined Franck-Condon and rotational constants analysis, J. Mol. Struct. **1223**, 129241 (2021).

[69] Q. Peng, Y. Niu, C. Deng, and Z. Shuai, Vibration correlation function formalism of radiative and non-radiative rates for complex molecules, Chem. Phys. **370**, 215 (2010).

[70] M. Biczysko, J. Bloino, and V. Barone, First principle simulation of vibrationally resolved $A^2 B_1 \leftarrow \tilde{X}^2 A_1$ electronic transition of phenyl radical, Chem. Phys. Lett. **471**, 143 (2009).

[71] W. Hua, K. Bennett, Y. Zhang, Y. Luo, and S. Mukamel, Study of double core hole excitations in molecules by x-ray double-quantum-coherence signals: A multi-configuration simulation, Chem. Sci. **7**, 5922 (2016).

[72] P. Macak, Y. Luo, and H. Ågren, Simulations of vibronic profiles in two-photon absorption, Chem. Phys. Lett. **330**, 447 (2000).

[73] Y. Zhang, W. Hua, K. Bennett, and S. Mukamel, Nonlinear spectroscopy of core and valence excitations using short x-ray pulses: Simulation challenges, Top. Curr. Chem. **368**, 273 (2016).

[74] F. Montorsi, F. Segatta, A. Nenov, S. Mukamel, and M. Garavelli, Soft X-ray spectroscopy simulations with multi-configurational wave function theory: Spectrum completeness, sub-eV accuracy, and quantitative reproduction of line shapes, J. Chem. Theory Comput. **18**, 1003 (2022).

[75] R. C. Couto, W. Hua, R. Lindblad, L. Kjellsson, S. L. Sorensen, M. Kubin, C. Bülow, M. Timm, V. Zamudio-Bayer, B. Von Issendorff, J. Söderström, J. T. Lau, J.-E. Rubensson, H. Ågren, and V. Carravetta, Breaking inversion symmetry by protonation: Experimental and theoretical NEXAFS study of the diazynium ion, N_2H^+, Phys. Chem. Chem. Phys. **23**, 17166 (2021).

[76] J. R. Reimers, A practical method for the use of curvilinear coordinates in calculations of normal-mode-projected displacements and Duschinsky rotation matrices for large molecules, J. Chem. Phys. **115**, 9103 (2001).